BRITISH JOURNAL OF
APPLIED PHYSICS

EDITOR

H. R. LANG

A.R.C.S., PH.D., F.INST.P.

Secretary of The Institute of Physics

VOLUME 5

including

Supplement No. 3

LONDON

THE INSTITUTE OF PHYSICS

BRITISH JOURNAL OF APPLIED PHYSICS

ADVISORY COMMITTEE, 1953–54

T. E. ALLIBONE, PH.D., D.SC., M.I.E.E., F.INST.P., F.R.S., *Chairman*

C. B. ALLSOPP, M.A., PH.D., D.SC., F.INST.P., *Representing The British Institute of Radiology*

H. BARRELL, A.R.C.S., D.SC., F.INST.P.

W. H. J. CHILDS, PH.D., D.SC., F.R.S.E., F.INST.P.

E. G. COX, D.SC., F.R.I.C., F.INST.P., F.R.S.

B. M. CROWTHER, M.A., PH.D., F.INST.P.

A. E. DE BARR, B.SC., F.INST.P., *Representing the X-ray Analysis Group*

B. P. DUDDING, M.B.E., A.R.C.S., PH.D., F.INST.P., *Honorary Secretary, The Institute of Physics*

J. S. FORREST, M.A., D.SC., M.I.E.E., F.INST.P., *Representing the Royal Meteorological Society*

E. K. FRANKL, M.A., *Representing the Stress Analysis Group*

A. G. GAYDON, A.R.C.S., PH.D., D.SC., A.INST.P., F.R.S., *Representing The Physical Society*

M. E. HAINE, M.SC., A.I.M.E.E., F.INST.P., *Representing the Electron Microscopy Group*

N. L. HARRIS, B.SC., F.INST.P.,

P. L. KIRBY, M.SC., A.INST.P., *Representing the North Eastern Branch*

H. KOLSKY, B.SC., PH.D., A.R.C.S., F.INST.P., *Representing The Physical Society*

F. LLEWELLYN JONES, M.A., D.PHIL., F.INST.P., *Representing the South Wales Branch*

W. J. MEREDITH, M.SC., F.INST.P.,

D. PARKINSON, PH.D., D.SC., F.INST.P., *Representing the Midland Branch*

F. Y. POYNTON, M.SC., F.INST.P., *Representing the Education Group*

J. S. RANKIN, B.SC., PH.D., M.I.MECH.E., F.INST.P., *Representing the Scottish Branch*

S. RODDA, B.SC., F.INST.P., *Representing the Electronics Group*

W. E. SCHALL, B.SC., F.INST.P., *Representing the Non-destructive Testing Group*

R. SCHNURMANN, M.SC., DR.RET.NAT., F.INST.P., *Representing the Manchester and District Branch*

E. W. H. SELWYN, A.R.C.S., B.SC., F.INST.P.

R. W. SILLARS, B.A., D.PHIL., A.M.I.E.E., F.INST.P.

E. C. STONER, B.A., PH.D., SC.D., F.INST.P., F.R.S.

C. SYKES, PH.D., D.SC., F.INST.P., F.R.S., *President, The Institute of Physics*

R. S. TEBBLE, B.SC., PH.D., A.INST.P., *Representing the Yorkshire Branch*

L. R. G. TRELOAR, PH.D., D.SC., F.INST.P., *Representing the London and Home Counties' Branch*

E. VAN SOMERON, B.SC., F.INST.P., *Representing the Industrial Spectroscopy Group*

S. WHITEHEAD, M.A., PH.D., D.SC., M.I.E.E., F.INST.P., *Honorary Treasurer, The Institute of Physics*

CONTENTS OF VOLUME 5

JANUARY 1954

SPECIAL ARTICLE — PAGE

Supersonic gas flow. By W. A. MAIR 1

ORIGINAL CONTRIBUTIONS

The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in a vacuum. By M. G. ROSSMANN and J. YARWOOD 7
Characteristics of radioluminescence in crystals. By G. T. WRIGHT and G. F. J. GARLICK 13
Line interaction in X-ray and electron diffraction photographs. By G. D. ARCHARD 19
The depression of liquid surfaces by gas jets. By R. D. COLLINS and H. LUBANSKA 22
The sources of electron-induced contamination in kinetic vacuum systems. By A. E. ENNOS 27
An improved experimental iteration method for use with resistance-network analogues. By G. LIEBMANN and R. BAILEY 32
The high-pressure glow discharge in air. By W. A. GAMBLING and H. EDELS 36
Approximate electrode shapes for a cylindrical electron beam. By E. R. HARRISON 40
The preparation and use of cells for the realization of the triple point of water. By C. R. BARBER, R. HANDLEY and E. F. G. HERINGTON 41

NOTES AND NEWS

New books 45
Notes and comments 48

FEBRUARY 1954

SPECIAL ARTICLE

Ionization processes in the electrical breakdown of gases. By F. LLEWELLYN JONES 49

ORIGINAL CONTRIBUTIONS

A standard light source of very low intensity based on the Čerenkov effect. By W. ANDERSON and E. H. BELCHER 53
Accuracy limitations of β-ray thickness measurement. By L. MANDEL 58
Evaporated carbon films for use in electron microscopy. By D. E. BRADLEY 65
High vacuum applications of polythene. By J. F. DUNCAN and D. T. WARREN 66
The effect of the finite light sources of measuring instruments on the determination of electron diffraction ring radii. By G. D. ARCHARD 69
A photometric method of determining configuration factors. By D. I. LAWSON and D. HIRD 72
The thermionic emission from thin films of barium and strontium oxide. By J. WOODS and D. A. WRIGHT . 74

NOTES AND NEWS

New books 77
Notes and comments 80

MARCH 1954

SPECIAL ARTICLES

Atomic energy—how it all began. By O. R. FRISCH . 81
The physical properties of gelatin solutions and gels. By A. G. WARD 85

ORIGINAL CONTRIBUTIONS — PAGE

Heat transfer by free convection in an open thermosyphon tube. By B. W. MARTIN and H. COHEN 91
An evaporated carbon replica technique for use with the electron microscope and its application to the study of photographic grains. By D. E. BRADLEY 96
The determination of the texture of sheet steel from torque curves. By L. R. BLAKE 99
Some observations on the time edge effect in Catalin 800. By A. R. MORRIS 104
The correction of an error in the determination of high resistances by capacitor discharge. By G. C. CURTIS 106
Some effects of slow electron bombardment in thermionic valves. By D. A. WRIGHT 108

NOTES AND NEWS

New books 112
Notes and comments 114

APRIL 1954

SPECIAL ARTICLES

New advances in the junction transistor. By E. W. HEROLD 115
The noise of valves. By C. S. BULL 127

ORIGINAL CONTRIBUTIONS

The evaluation of interferograms by displacement and stereoscopic methods. By J. W. GATES 133
The mechanism of the resolution of water-in-oil emulsions by electrical treatment. By C. A. R. PEARCE . . 136
A note on an improved planimetric method of harmonic analysis. By J. CREASE and M. J. TUCKER . . 143
An application of conformal transformation to the investigation of the magnetic field between galvanometer pole-pieces. By W. SNOWDON and N. DAVY . 146
A magnetic method for the determination of γ-loops in binary iron alloys and its application to the iron-silicon system. By J. CRANGLE 151

NOTES AND NEWS

New books 155
Notes and comments 156

MAY 1954

SPECIAL ARTICLE

The acceleration of heavy ions in a fixed-frequency cyclotron. By D. WALKER, J. H. FREMLIN, W. T. LINK and K. G. STEPHENS 157

CONFERENCE REPORT

Summarized proceedings of a conference on electron microscopy—London, November 1953. By C. E. CHALLICE 165

ORIGINAL CONTRIBUTIONS

Some measurements of the relative dielectric strength of gases. By N. R. MCCORMICK and J. D. CRAGGS . 171
Control systems with quasi-critical damping. By A. T. FULLER 174
Potential distribution of symmetrical cylindrical electron lenses. By G. D. ARCHARD 179

PAGE

The emissivity of a groove. By L. F Daws . . . 182
The rolling resistance of a wheel with a solid rubber tyre. By I. Evans 187
The growth of the liquid bridge in an electrical contact. By P. M. Davidson 189
Automatic electron trajectory plotting using the electrolytic tank analogue. By B. O. Baker 191

NOTES AND NEWS

Correspondence:

A simple bridge method of measurement of the time constant of exponential decays. From D. F. Gibbs; C. M. Burrell 196
The measurement of spectral distributions of luminescent materials above room temperature. From A. Langridge and M. P. Lord 196

New books 197

Notes and comments 198

JUNE 1954

SPECIAL ARTICLE

Some thoughts on the future of natural science, with illustrations from the growth of X-ray diffraction work in the United States. By R. W. G. Wyckoff . . 199

ORIGINAL CONTRIBUTIONS

The frequency response of a certain class of non-linear feedback systems. By J. C. West and J. L. Douce . . 204
Spectrophotometry of fluorescent pigments. By R. Donaldson 210
Direct-reading metal spectroscopy with a d.c. arc. By S. C. Baker 215
The specific heat of polymerized Araldite and Wood's metal between 1·5 and 20° K. By D. H. Parkinson and J. E. Quarrington 219
Densitometric evaluation of coal-dust stains on filter paper. By J. G. Dawes 221
Transverse oscillations in travelling strings. By R. A. Sack 224
A note on the selection of observers for primary standard optical pyrometry. By M. C. Probine and S. Bertaud 227

NOTES AND NEWS

New books 229

Notes and comments 230

JULY 1954

SPECIAL ARTICLE

Diffuse X-ray scattering and the physical properties of crystals. By W. A. Wooster 231

ORIGINAL CONTRIBUTIONS

The use and scope of Iridium 192 for the radiography of steel. By R. Halmshaw 238
The critical evaluation of compression data for liquids and a revision of the isotherms of mercury. By K. E. Bett, K. E. Weale and D. M. Newitt 243
On the application of Ångström's method of measuring thermal conductivity. By C. H. Bosanquet and R. Aris 252
The anode fall in a glow discharge. By C. J. F. Chaundy 255

PAGE

The determination of grain size from spotty X-ray diffraction rings. By P. B. Hirsch 257
The flux linkage with a search coil produced by a coaxial uniformly magnetized prolate spheroid. By G. H. Hunt 260
Some consideration of the errors of brightness and two-colour types of spectral radiation pyrometer. By E. C. Pyatt 264

NOTES AND NEWS

Correspondence:

The molecular flow of gases. From D. W. Stops . . 269
A check on the standard observer data at 4358 Å. From W. Harrison 269
Line interaction in X-ray and electron diffraction photographs. From D. M. C. MacEwan; G. D. Archard 270
The fitting of pyrophyllite for minimum distortion. From V. E. Dyke 270
The noise of valves. From C. S. Bull 270

New books 271

Notes and comments 272

AUGUST 1954

SPECIAL ARTICLE

Statistical concepts in theoretical physics. By R. Fürth 273

ORIGINAL CONTRIBUTIONS

Sub-normal discharges. By K. G. Emeleus and J. A. Burns 277
Triaxial stresses caused by notches. By A. F. C. Brown 280
The effect of latent heat on numerical solutions of the heat flow equation. By P. H. Price and M. R. Slack . 285
The β-ray absorption spectrum of $^{147}Pm_{61}$ and its application to thickness measurement. By L. Mandel . 287
A method of investigating the secondary emissive properties of insulators. By C. N. W. Litting . . 289
Requirements contributing to the design of devices used in correcting electron lenses. By G. D. Archard . 294

NOTES AND NEWS

New books 300

Notes and comments 304

SEPTEMBER 1954

SPECIAL ARTICLE

The thermodynamic approach to solid structure. By V. Daniel 305

ORIGINAL CONTRIBUTIONS

Methods for studying the thermal resistances of sprayed and electro-deposited metal coatings. By R. W. Powell and M. J. Hickman 312
The capacity and field of a split cylindrical condenser, using the method of inversion. By H. J. Peake and N. Davy 316
X-ray determination of crystallinity in deformed natural rubber. By S. C. Nyburg 321
An improved technique for the micro-electrophoresis of oil drops. By D. A. Smith 325
The elliptic surface wave. By A. E. Karbowiak . . 328
The electrical breakdown of pre-stressed air at atmospheric pressure. By D. R. Hardy and H Wroe . . 335

NOTES AND NEWS — PAGE

Correspondence:

Reflectance measurements on siliceous dusts and powders. From E. V. SMITH 340

Application of a tracer to cathode-gettering and gas-adsorption problems. From F. DE BOER and W. F. NIKLAS 341

New books 342

Notes and comments 344

OCTOBER 1954

SPECIAL ARTICLE

Summarized proceedings of a conference on X-ray analysis—Cambridge, April 1954. By W. COCHRAN . 345

ORIGINAL CONTRIBUTIONS

A method of examining selected areas of surfaces using replicas and the electron microscope. By G. R. BOOKER 349

A potentiostat for corrosion study. By M. H. ROBERTS 351

The glass sealing properties of titanium and zirconium. By H. RAWSON and E. P. DENTON 352

Load-deflexion relations of rubber bush mountings. By J. E. ADKINS and A. N. GENT 354

The experimental determination of the speed of a vacuum pump and of components of a vacuum system. By C. W. OATLEY 358

Resistance-network analogues with unequal meshes or subdivided meshes. By G. LIEBMANN 362

The calibration of circular scales and precision polygons. By A. H. COOK 367

The capacity and field of a split cylindrical condenser when the conductors differ in length. By H. J. PEAKE and N. DAVY 371

NOTES AND NEWS

New books 374

Notes and comments 376

NOVEMBER 1954

SPECIAL ARTICLE

The near infra-red absorption spectra of natural and synthetic fibres. By A. ELLIOTT, W. E. HANBY and B. R. MALCOLM 377

ORIGINAL CONTRIBUTIONS

Further tests on the stability of analytical weights in chemical laboratories. By P. H. BIGG and F. H. BURCH 382

The use of semi-conductors in thermoelectric refrigeration. By H. J. GOLDSMID and R. W. DOUGLAS . . . 386

The emission from hot cathodes in gas discharges. By A. E. PENGELLY and D. A. WRIGHT 391

Some properties of symmetrical slit ("cylindrical") electron lenses. By G. D. ARCHARD 395

A theoretical investigation of the temperature distribution in the metal cutting process. By A. C. RAPIER . . 400

The two-dimensional electric field of a curved-sided wall or groove on an infinite plane. By N. H. LANGTON and N. DAVY 405

NOTES AND NEWS

Correspondence:

The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in vacuum. From L. HOLLAND; J. YARWOOD 411

Note on the resistance-network analogue solution of field problems of spherical symmetry. From G. LIEBMANN 412

New books 413

Notes and comments 416

DECEMBER 1954

CONFERENCE REPORT

Summarized proceedings of a conference on X-ray analysis—London, November 1953. By U. W. ARNDT and J. B. NELSON 417

ORIGINAL CONTRIBUTIONS

Determination of thermal conductivities at high temperatures. By K. S. KRISHNAN and S. C. JAIN . . . 426

The two-dimensional magnetic or electric field of a single isolated pole-piece. By N. H. LANGTON and N. DAVY 431

The mechanical impedance of the human mastoid process. By R. S. DADSON, D. W. ROBINSON and R. G. P. GREIG 435

The shape of the emission bands of luminescent solids. By C. C. VLAM 443

High current spark channels. By J. E. ALLEN and J. D. CRAGGS 446

NOTES AND NEWS

Correspondence:

The correction of an error in the determination of high resistances by capacitor discharges. From D. K. MCCLEERY; G. C. CURTIS 454

The solarization of polymethylmethacrylate. From K. J. MCCREE 454

The effect of intensifying screens on X-ray line breadths. From R. I. GARROD and J. H. AULD 454

New books 455

Notes and comments 458

SUPPLEMENT No. 3

THE PHYSICS OF PARTICLE SIZE ANALYSIS

SESSION A. RELATIVE MOTION OF PARTICLES AND FLUIDS: SIZE SEPARATION

Paper — *Page*

Survey of the relative motion of particles and fluids. By P. G. W. HAWKSLEY, British Coal Utilisation Research Association . . . . . S 1

A.1. The effective viscosity of suspensions. By Prof. G. J. KYNCH . . . . . S 5

A.2. The terminal velocity and structure of airborne aggregates of microscopic spheres. By V. TIMBRELL . . . . . S 12

A.3. Dust shadows below a cylinder containing a suction orifice and deposition of particles upon the cylinder. By C. N. DAVIES, and V. PEETZ . . S 17

A.4. A comparison of methods for particle size analysis. By B. A. JARRETT, and H. HEYWOOD . . . S 21

A.5. Theory of size classification of airborne clouds by elutriation. By W. H. WALTON . . . . S 29

General discussion on Session A . . . . . S 39

SESSION B. RELATIVE MOTION OF PARTICLES AND FLUIDS: MOLECULAR PHENOMENA

Paper — *Page*

B.1. The measurement of surface area of finely divided solids by the flow of gases through them. By R. M. BARRER . . . . . S 41

B.2. An examination of the behaviour of adsorbed molecules by the flow of gases through microporous media. By R. M. BARRER . . . . . S 49

B.3. The flow of gases through packed beds of colloidal carbon blacks. By H. W. DAVIDSON, and B. M. TURK . . . . . S 55

B.4. Electroviscosity in dilute heterodisperse suspensions. By G. A. H. ELTON and F. G. HIRSCHLER . S 60

SESSION C. SCATTERING AND ADSORPTION OF LIGHT BY PARTICLES

Survey of scattering and absorption of light by particles. By C. N. DAVIES . . . . . S 64

C.1. Light scattering by polydisperse dust clouds. By J. MCK. ELLISON . . . . . S 66

C.2. Photoextinction measurements on spherical particles. By P. C. LEWIS, and G. F. LOTHIAN . . . S 71

General discussion on Session C . . . . . S 74

SESSION D. PARTICLE SHAPE FACTORS

D.1. The shape of ground petroleum coke particles. By T. BEIRNE and J. M. HUTCHEON . . . . S 76

D.2. The significance and application of shape factors in particle size analysis. By W. H. M. ROBINS . S 82

D.3. The terminal velocity and size of airborne dust particles. By V. TIMBRELL . . . . . S 86

D.4. The relation between free-falling speed and particle size of airborne dusts. By R. J. HAMILTON . . S 90

General discussion on Session D . . . . . S 95

SESSION E. VISUAL COUNTING AND SIZING OF MICROSCOPIC PARTICLES

E.1. Inter-laboratory checks on the counting of coal dust particles on thermal precipitator slides. By J. F. HOLDSWORTH, F. HENLEY PRICE and R. C. TOMLINSON . . . . . S 96

E.2. Factors in the design of a microscope eyepiece graticule for routine dust counts. By R. J. HAMILTON, J. F. HOLDSWORTH, and W. H. WALTON . . . . . S 101

E.3. A particle-profile test strip for assessing the accuracy of sizing irregularly shaped particles with a microscope. By H. H. WATSON and D. F. MULFORD . S 105

E.4. The electron microscopy of airborne dusts. By J. CARTWRIGHT . . . . . S 109

General discussion on Session E . . . . . S 120

SESSION F. AUTOMATIZED COUNTING AND SIZING: THEORY

Paper — *Page*

Survey of the automatic counting and sizing of particles. By W. H. WALTON . . . . S 121

F.1 Theory of particle sizing and counting by track scanning. By P. G. W. HAWKSLEY . . . . S 125

F.2 Some fundamental aspects of particle counting and sizing by line scans. By J. W. PHILLIPS . . S 133

SESSION G. AUTOMATIZED COUNTING AND SIZING: PHOTOELECTRONIC MACHINES

G.1. A photoelectronic instrument for counting and sizing aerosol particles. By F. T. GUCKER, JR., and D. G. ROSE . . . . . S 138

G.2. The automatic size analysis of dust deposits by means of an illuminated slit. By L. LE BOUFFANT and J. L. SOULÉ . . . . . S 143

G.3. The automatic counting of red blood cells. By E. H. COOKE-YARBOROUGH and R. E. WHYARD . . S 147

G.4. Stages in the development of an arrested scan type microscopic particle counter. By H. A. DELL . S 156

G.5. Testing a counting machine. By R. J. COURSHEE . S 161

G.6. The design and construction of a photoelectronic scanning machine for sizing microscopic particles. By P. G. W. HAWKSLEY, J. H. BLACKETT, E. W. MEYER and A. E. FITZSIMMONS . . . S 165

G.7. An automatic system for obtaining particle size distributions with the aid of the flying spot microscope. By W. K. TAYLOR . . . . S 173

General discussion on Sessions F and G . . S 175

SESSION H. GENERAL

H.1. The application of methods of particle size and surface area measurement to whiting. By G. E. BESSEY and D. C. SOUL . . . . . S 181

H.2. A microscope method of measuring thicknesses of fine particles. By W. H. M. ROBINS . . S 189

H.3. Some methods of dispersing difficult powders. By M. G. HARWOOD . . . . . S 193

H.4. The adhesion of dust particles. By D. W. JORDAN S 194

H.5. The aggregation of aerosols. By J. M. DALLAVALLE, CLYDE ORR, JR, and B. L. HINKLE, The Georgia Institute of Technology . . . . . S 198

General discussion on Session H . . . . S 206

SUMMARY OF CONFERENCE AND CONCLUDING REMARKS

By H. HEYWOOD . . . . . S 209

THE DEMONSTRATIONS . . . . . S 210

SUBJECT INDEX TO THE PHYSICS OF PARTICLE SIZE ANALYSIS . . . . . S 211

NAME INDEX TO THE PHYSICS OF PARTICLE SIZE ANALYSIS . . . . . S 215

SUBJECT INDEX TO VOLUME 5 459

AUTHOR INDEX TO VOLUME 5 464

INDEX TO VOLUME 5

SUBJECT INDEX

Abstracts, Analytical 114
Adhesion, Film structure and 344
Adsorption, gas-, problems, Application of a tracer to cathode-gettering and 341
Air
  pre-stressed, The electrical breakdown of, at atmospheric pressure 335
  The high-pressure glow discharge in 36
Analogue
  electrolytic tank, Automatic electron trajectory plotting using the 191
  resistance-network, solution of field problems of spherical symmetry, Note on the 412
Analogues, resistance-network, An improved experimental iteration method for use with 32
Analysis, harmonic, A note on an improved planimetric method of 143
Ångström's method of measuring thermal conductivity, On the application of 252
Anode fall in a glow discharge, The 255
Araldite, polymerized, The specific heat of, and Wood's metal between 1·5 and 20° K 219
Areas, selected, of surfaces, A method of examining, using replicas and the electron microscope 349
Atomic energy—how it all began 81

Barium and strontium oxide, The thermionic emission from thin films of 74
Beta-ray
  absorption spectrum of $^{147}Pm_{,1}$ and its application to thickness measurement, The 287
  thickness measurement, Accuracy limitations of 58
Books, New
  Advanced treatise on physical chemistry, An. Vol. 4 79
  Advances in electronics, Vol. 5 375
  Amplification and distribution of sound, The (3rd edition), 456
  Amplitude-frequency characteristics of ladder networks 457
  Applied elasticity 302
  Atomization of liquid fuels, The 112
  Atomphysik, Vols. III and IV 456
  Atmospheric electricity 112
  Automatic digital calculators 46
  Bessel functions and formulae 197
  Communication theory 113
  Complex variable theory and transform calculus (2nd edition) 197
  Crystal growth and dislocations 79
  Determination of crystal structures, The 302
  Dielectric aerials 46
  Dimensional methods and their applications 155
  Dipole moments 113
  Dislocations and plastic flow in crystals 47
  Elasticity, plasticity and structure of matter 415
  Electrical breakdown of gases, The 300
  Electron diffraction 45
  Electron optics 46
  Elementary introduction to molecular spectra 342
  Experimental nuclear physics, Vol. 1 45, Vol. 2 229
  Flames: Their structure, radiation and temperature 46
  Flow properties of disperse systems 343
  Fluorescence of solutions 78
  Flying saucers 271
Books, New (*continued*)
  Foodstuffs, their plasticity, fluidity and consistency 112
  Fundamental processes of electrical contact phenomena 47
  Geometrical mechanics and de Broglie waves 456
  Grössengleichungen, einheiten und dimensionen 77
  Heat conduction, with engineering, geological and other applications 455
  High altitude rocket research 300
  High voltage laboratory technique 413
  Industrial electronics 374
  Introduction to a study of mechanical vibration 155
  Introduction to electron microscopy 113
  Introduction to relaxation methods, An 112
  Introduction to solid state physics 77
  Introduction to tensors, spinors and relativistic wave equations (relative structure) 342
  Introduction to the theory of seismology, An 302
  Investigations of the alkali-content of refractories using the Riehm–Lange flame photometer 457
  Low frequency amplification 229
  Low temperature physics 343
  Measurement of humidity 78
  Measurement techniques in mechanical engineering 155
  Mechanical properties of wood and paper 77
  Memorandum on gamma-ray sources for radiography 413
  Metallurgy of the rare metals. No. 1, *Chromium. No.* 2, *Zirconium* 301
  Methods of algebraic geometry. Vol. 3. *Birational geometry* 374
  Methods of theoretical physics (Parts 1 and 2) 197
  Microwave spectroscopy 78, 229
  Millimicrosecond pulse techniques 375
  Monographies C, Physique Mathematique: 11. *Les fonctions orthogonales dans les problems aux limites de la physique mathematique* 271
  Nuclear moments 413
  Physics and applications of secondary electron emission, The 414
  Present state of physics, The 415
  Principles of transistor circuits 303
  Printing of mathematics, The 375
  Probability and information theory, with applications to radar 112
  Problems in the advanced theory of functions. Vol. 2 47
  Proceedings of the first international congress of electron microscopy, Memoire hors série no. 1 413
  Proceedings of the second international congress on rheology, Oxford, 26–31 *July*, 1953 456
  Proceedings of the Western Computer Conference 78
  Progress in biophysics and biophysical chemistry, Vol. 4 455
  Progress in cosmic ray physics, Vol. 2 301
  Progress in nuclear physics, Vol. 3 197
  Properties of metallic surfaces 45
  Radioactive isotopes 47
  Radio engineering, Vol. 1 155
  Rapid analytical method for the determination of alumina, silica and iron oxide in refractories by the photometry of organic complex reagents, A 457
  Reihenentwicklungen in der mathematischen Physik 271
  Relativity—the special and the general theory 303
  Reports on progress in physics. Vol. XVI 342 Vol. XVII 457
  Statistical approach to X-ray structure analysis, The 414

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

Books, New (*continued*)
Statistical astronomy 155
Structure and properties of solid surfaces 457
Structure of metals and alloys, The 415
Structure reports for 1945–46, Vol. 10 303
Styrene 79
Sun, the solar system, The, Vol. 1 414
Technical aspects of sound 78
Technische hydro-und aeromechanik 457
Television receiver design, Monograph 2. *Flywheel synchronization of saw-tooth generators* 300
Temperature measurement in engineering, Vol. 1 374
Theory of metals (2nd edition), *The* 271
Thermionic valves, their theory and design 229
Twinning and diffusionless transformations in metals 374
Vacuum tube oscillators 47
Wärmespannungen infolge stationärer Temperaturfelder 414
Zirkonium, seine Herstellung Ergenschaften und Anwendungen in der Vakuumtechnik 415
Bowen Prizes 304
Breakdown
electrical, of gases, Ionization processes in the 49
The electrical, of pre-stressed air at atmospheric pressure 335
Bridge
, liquid, The growth of the, in an electrical contact 189
method of measurement of the time constant of exponential decays, A simple 196
Brightness and two-colour types of spectral radiation pyrometer, Some consideration of the errors of 264
British Standards Institution, The 416

Capacitor discharges, The correction of an error in the determination of high resistances by 454
Capacity
and field of a split cylindrical condenser, using the method of inversion, The 316
and field of a split cylindrical condenser when the conductors differ in length, The 371
Carbon
crucibles, The use of, in measurements on the rate of evaporation of liquid metals in a vacuum 7, 411
films, Evaporated, for use in electron microscopy 65
replica technique, evaporated, for use with the electron microscope and its application to the study of photographic grains 96
Catalin 800, Some observations on the time edge effect in 104
Cathodes
hot, The emission from, in gas discharges 391
oxide coated, International convention on 156
Cells, The preparation and use of, for the realization of the triple point of water 41
Čerenkov effect, A standard light source of very low intensity based on the 53
Channels, High current spark 446
Coal-dust stains on filter paper, Densitometric evaluation of 221, 376
Coal, Proceedings of a Colloquium on the Hydraulic Transport of 304, 344
Coatings, sprayed and electro-deposited metal, Methods for studying the thermal resistances of 312
Coil, search, The flux linkage with a, produced by a coaxial uniformly magnetized prolate spheroid 260
Commission, International, Recommendations of the, on Radiological Units 156
Communications and Electronics 376
Competition, Bicentenary 48
Components
electronic, Improvement of 80
of servo-mechanisms, Terminology of 304
Compression data for liquids, The critical evaluation of, and a revision of the isotherms of mercury 243
Computers, small digital, Design and application of 304
Condenser
split cylindrical, The capacity and field of a, using the method of inversion 316
split cylindrical, The capacity and field of a, when the conductors differ in length 371
Conductivities, thermal, Determination of, at high temperatures 426
Conductivity, thermal, On the application of Ångström's method of measuring 252
Conference
, International, on semiconductors 80
on meteorology—Toronto, 1953 458
on X-ray analysis—London, November 1953, Summarized proceedings of a 417
, Second Radioisotope 48
Configuration factors, A photometric method of determining 72
Conformal transformation, An application of, to the investigation of the magnetic field between galvanometer pole-pieces 146
Contact, electrical, The growth of the liquid bridge in an 189
Contamination, electron-induced, The sources of, in kinetic vacuum systems 27
Control systems with quasi-critical damping 174
Convection, free, Heat transfer by, in an open thermosyphon tube 91
Convention, International, on oxide coated cathodes 156
Corrosion study, A potentiostat for 351
Crucibles, carbon, The use of, in measurements on the rate of evaporation of liquid metals in a vacuum 7, 411
Crystal data and numerical structure factor tables 80
Crystallinity in deformed natural rubber, X-ray determination of 321
Crystallography, Third general assembly and international congress of 198
Crystals
Characteristics of radioluminescence in 13
Diffuse X-ray scattering and the physical properties of 231
Cyclotron, fixed-frequency, The acceleration of heavy ions in a 157

Damping, quasi-critical, Control systems with 174
Densitometric evaluation of coal-dust stains on filter paper 221, 376
Depression of liquid surfaces by gas jets, The 22
Dielectric strength of gases, Some measurements of the relative 171
Diffraction
photographs, electron, Line interaction in X-ray and 270
ring radii, electron, The effect of the finite light sources of measuring instruments on the determination of 69
rings, spotty X-ray, The determination of grain size from 257
, X-ray and electron, photographs, Line interaction in 19
work, X-ray, in the United States, Some thoughts on the future of natural science, with illustrations from the growth of 199
Discharge
capacitor, The correction of an error in the determination of high resistances by 106
glow, The anode fall in a 255
, The high-pressure glow, in air 36
Discharges
capacitor, The correction of an error in the determination of high resistances by 454
gas, The emission from hot cathodes in 391
Sub-normal 277
Displacement and stereoscopic methods, The evaluation of interferograms by 133
Dusts, siliceous, and powders, Reflectance measurements on 340

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

Electrode shapes for a cylindrical electron beam, Approximate 40
Electron
beam, cylindrical, Approximate electrode shapes for a 40
bombardment in thermionic valves, Some effects of slow 108
diffraction photographs, Line interaction in X-ray and 19, 270
diffraction ring radii, The effect of the finite light sources of measuring instruments on the determination of 69
lenses, Requirements contributing to the design of devices used in correcting 294
lenses, symmetrical cylindrical, Potential distribution of 179
lenses, symmetrical slit ("cylindrical"), Some properties of 395
microscope, An evaporated carbon replica technique for use with the, and its application to the study of photographic grains 96
microscopy, Evaporated carbon films for use in 65
microscopy, Summarized proceedings of a conference on—London, November 1953 165
trajectory plotting, Automatic, using the electrolytic tank analogue 191
Electronic components, Improvement of 80
Elliptic surface wave, The 328
Emission
bands of luminescent solids, The shape of the 443
from hot cathodes in gas discharges, The 391
thermionic, from thin films of barium and strontium oxide, The 74
Emissive properties, secondary, of insulators, A method of investigating the 289
Emissivity of a groove, The 182
Emulsions, water-in-oil, The mechanism of the resolution of, by electrical treatment 136
Equation, non-linear differential, Solution of a 48
Evaporation
of liquid metals in a vacuum, The use of carbon crucibles in measurements on the rate of 7, 411
Exhibition, The Physical Society's annual, of instruments and apparatus 80

Fibres, natural and synthetic, The near infra-red absorption spectra of 377
Field
of a split cylindrical condenser, The capacity and, using the method of inversion 316
of a split cylindrical condenser when the conductors differ in length, The capacity and 371
problems of spherical symmetry, Note on the resistance-network analogue solution of 412
, The two-dimensional electric field of a curved-sided wall or groove on an infinite plane 405
, The two-dimensional magnetic or electric, of a single isolated pole-piece 431
Film structure and adhesion 344
Flow
of gases, The molecular 269
, Supersonic gas 1
Fluorescent pigments, Spectrophotometry of 210
Flux linkage with a search coil produced by a coaxial uniformly magnetized prolate spheroid, The 260
Frequency response of a certain class of non-linear feedback systems, The 204

Gamma-loops in binary iron alloys, A magnetic method for the determination of, and its application to the iron-silicon system 151
Gas
discharges, The emission from hot cathodes in 391
flow, Supersonic 1
jets, The depression of liquid surfaces by 22
Gases,
electrical breakdown of, Ionization processes in the 49
Some measurements of the relative dielectric strength of 171
The molecular flow of 269
Gels, The physical properties of gelatin solutions and 85
Gettering, cathode-, and gas-adsorption problems, Application of a tracer to 341
Glow discharge in air, The high-pressure 36
Grain size, The determination of, from spotty X-ray diffraction rings 257
Groove, The emissivity of a 182

Harmonic analysis, A note on an improved planimetric method of 143
Hazards, Radiation, in industry 376
Heat
flow equation, The effect of latent heat on numerical solutions of the 285
transfer by free convection in an open thermosyphon tube 91
Hydraulic Transport of Coal, Proceedings of a Colloquium on the 304, 344

Impedance, The mechanical, of the human mastoid process 435
Index, Ten Year, of Proceedings of The Institution of Electrical Engineers 230
Infra-red absorption spectra, The near, of natural and synthetic fibres 377
Institute of Physics, The 272
Elections to 80, 156, 230, 304, 416, 458
Insulators, A method of investigating the secondary emissive properties of 289
Interferograms, The evaluation of, by displacement and stereoscopic methods 133
Ionization processes in the electrical breakdown of gases 49
Ions, heavy, The acceleration of, in a fixed-frequency cyclotron 157
Iridium 192 for the radiography of steel, The use and scope of 238
Iron alloys, binary, A magnetic method for the determination of gamma-loops in, and its application to the iron-silicon system 151
Isotherms of mercury, The critical evaluation of compression data for liquids and a revision of the 243

Jets, gas, The depression of liquid surfaces by 22

Laboratory and Workshop Notes 1950–1952 416
Latent heat, The effect of, on numerical solutions of the heat flow equation 285
Lenses
electron, Requirements contributing to the design of devices used in correcting 294
symmetrical cylindrical electron, Potential distribution of 179
Letter Symbols, Signs and Abbreviations. Part 1 *General*, B.S. 1991: 1954 198
Liquid
bridge, The growth of the, in an electrical contact 189
metals, evaporation of, in a vacuum, The use of carbon crucibles in measurements on the rate of 7
surfaces, The depression of, by gas jets 22
Liquids, The critical evaluation of compression data for, and a revision of the isotherms of mercury 243
Load-deflexion relations of rubber bush mountings 354
Luminescent
materials above room temperature, The measurement of spectral distributions of 196
solids, The shape of the emission bands of 443

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

Magnetic
field between galvanometer pole-pieces, An application of conformal transformation to the investigation of the 146
method for the determination of gamma-loops in binary iron alloys and its application to the iron-silicon system, A 151
Magnetized prolate spheroid, coaxial uniformly, The flux linkage with a search coil produced by a 260
Mastoid process, human, The mechanical impedance of the 435
Mercury, the isotherms of, The critical evaluation of compression data for liquids and a revision of 243
Meshes, unequal, or subdivided meshes, Resistance-network analogues with 362
Metal
cutting process, A theoretical investigation of the temperature distribution in the 400
spectroscopy, Direct-reading, with a d.c. arc 215
Metals
liquid, in vacuum, The use of carbon crucibles in measurements on the rate of evaporation of 411
, The mechanism of phase transformations in 344
Meteorology, Recent advances in,—Toronto, 1953 458
Micro-electrophoresis of oil drops, An improved technique for the 325
Miscroscope, electron, An evaporated carbon replica technique for use with the, and its application to the study of photographic grains 96
Microscopy, electron
Evaporated carbon films for use in 65
Summarized proceedings of a conference on,—London, November 1953 165
Mountings, rubber bush, Load-deflexion relations of 354

Natural science, Some thoughts on the future of, with illustrations from the growth of X-ray diffraction work in the United States 199
Noise, extra-terrestrial radio, Report on discrete sources of 230
Notes on the preparation of contributions to the Institute's Journals and other publications 376

Observer data, standard, A check on the, at 4358 Å 269
Observers for primary standard optical pyrometry, A note on the selection of 227
Oil drops, An improved technique for the micro-electrophoresis of 325
Optica Acta 198
Oscillations, Transverse, in travelling strings 224

Petroleum Measurement Tables, ASTM-IP 48
Phase transformations in metals, The mechanism of 344
Photographic grains, An evaporated carbon replica technique for use with the electron microscope and its application to the study of 96
Physical Society's annual exhibition of instruments and apparatus, The 80
Planimetric method of harmonic analysis, A note on an improved 143
Pole-piece, single isolated, The two-dimensional magnetic or electric field of a 431
Polygons, precision, The calibration of circular scales and 367
Polymethylmethacrylate, The solarization of 454
Polythene, High vacuum applications of 66
Potential distribution of symmetrical cylindrical electron lenses 179
Potentiostat for corrosion study, A 351
Powders, Reflectance measurements on siliceous dusts and 340
Prize, Napier Shaw Essay 198
Prizes for original contributions published in the *Journal of Scientific Instruments* 304
Proceedings of The Institution of Electrical Engineers, Ten Year Index of 230
Pump, vacuum, The experimental determination of the speed of a, and of components of a vacuum system 358
Pyrometer, spectral radiation, Some consideration of the errors of brightness and two-colour types of 264
Pyrometry, primary standard optical, A note on the selection of observers for 227
Pyrophyllite, The firing of, for minimum distortion 270

Radiation hazards in industry 376
Radiation Research 230
Radiography of steel, The use and scope of Iridium 192 for the 238
Radioisotope Conference, Second 48
Radiological Units, Recommendations of the International Commission on 156
Radioluminescence in crystals, Characteristics of 13
Radio noise, extra-terrestrial, Report on discrete sources of 230
Recording, sound, Methods of 48
Reflectance measurements on siliceous dusts and powders 340
Refrigeration, thermoelectric, The use of semiconductors in 386, 458
Register of British manufacturers 80
Resistance
-network analogue solution of field problems of spherical symmetry, Note on the 412
-network analogues with unequal meshes or subdivided meshes 362
, rolling, of a wheel with a solid rubber tyre, The 187
Resistances
high, The correction of an error in the determination of, by capacitor discharge 106, 454
thermal, of sprayed and electro-deposited metal coatings, Methods for studying the 312
Rubber
bush mountings, Load-deflexion relations of 354
, deformed natural, X-ray determination of crystallinity in 321

Scales, circular, and precision polygons, The calibration of 367
Scattering, Diffuse X-ray and the physical properties of crystals 231
Scientia Electrica 344
Sealing properties, The glass, of titanium and zirconium 352
Semiconductors
International conference on 80
The use of, in thermoelectric refrigeration 386, 458
Servomechanisms
and related subjects, Bibliography of information on 114
Terminology of components of 304
Solarization of polymethylmethacrylate, The 454
Solid structure, The thermodynamic approach to 305
Solutions, gelatin, and gels, The physical properties of 85
Sound, Methods of recording 48
Source, A standard light, of very low intensity based on the Čerenkov effect 53
Sources, The effect of the finite light, of measuring instruments on the determination of electron diffraction ring radii 69
Spark channels, High current 446
Specific heat of polymerized Araldite and Wood's metal between 1·5 and 20° K, The 219
Spectra, The near infra-red absorption, of natural and synthetic fibres 377
Spectral distributions of luminescent materials above room temperature, The measurement of 196
Spectrometry, Mass, Second annual meeting of ASTM Committee on 156
Spectrophotometry of fluorescent pigments 210

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

Spectroscopy
Colloquium, Fifth International 114
, Direct-reading metal, with a d.c. arc 215
Spectrum, beta-ray absorption, of $^{147}Pm_{51}$ and its application to thickness measurement, The 287
Speed of a vacuum pump and of components of a vacuum system, The experimental determination of the 358
Spheroid, coaxial uniformly magnetized prolate, The flux linkage with a search coil produced by a 260
Stains, coal-dust, on filter paper, Densitometric evaluation of 221, 376
Standards Institution, The British 416
Statistical concepts in theoretical physics 273
Steel
sheet, The determination of the texture of, from torque curves 99
, the radiography of, The use and scope of Iridium 192 for 238
Stereoscopic methods, The evaluation of interferograms by displacement and 133
Stresses, Triaxial, caused by notches 280
Strings, travelling, Transverse oscillations in 224
Strontium oxide, The thermionic emission from thin films of barium and 74
Supersonic gas flow 1
Surface wave, The elliptic 328
Surfaces
A method of examining selected areas of, using replicas and the electron microscope 349
liquid, The depression of, by gas jets 22
Symposium, High vacuum 156

Tables
ASTM-IP Petroleum Measurement 48
Crystal data and numerical structure factor 80
Temperature distribution, A theoretical investigation of the, in the metal cutting process 400
Temperatures, high, Determination of thermal conductivities at 426
Texture of sheet steel, The determination of the, from torque curves 99
Theoretical physics, Statistical concepts in 273
Thermal conductivity, On the application of Ångström's method of measuring 252
Thermionic
emission from thin films of barium and strontium oxide, The 74
valve, The jubilee of the invention of the 416
valves, Educational booklets on 114
valves, Some effects of slow electron bombardment in 108
Thermodynamic approach to solid structure, The 305
Thermoelectric refrigeration, The use of semiconductors in 386
Thermosyphon tube, Heat transfer by free convection in an open 91
Time
constant of exponential decays, A simple bridge method of measurement of the 196
edge effect in Catalin 800, Some observations on the 104
Tin, The properties of 344
Titanium and zirconium, The glass sealing properties of 352
Tracer, Application of a, to cathode-gettering and gas-adsorption problems 341
Transistor, junction, New advances in the 115
Triple point of water, The preparation and use of cells for the realization of the 41

Units, Radiological, Recommendations of the International Commission on 156

Vacuum
High-applications of polythene, 66
, High, symposium, 156
pump, The experimental determination of the speed of a, and of components of a vacuum system 358
systems, kinetic, The sources of electron-induced contamination in 27
Valve, thermionic, The jubilee of the invention of the 416
Valves
The noise of 127, 270
thermionic, Educational booklets on 114
thermionic, Some effects of slow electron bombardment in 108

Water, the triple point of, The preparation and use of cells for the realization of 41
Wave, The elliptic surface 328
Weights, analytical, Further tests on the stability of, in chemical laboratories 382
Welding polythene and polyvinyl chloride 376
Wheel with a solid rubber tyre, The rolling resistance of a 187
Wood's metal, The specific heat of polymerized Araldite and, between 1·5 and 20° K 219

X-ray
analysis, Summarized proceedings of a conference on,—Cambridge, April 1954 345
analysis, Summarized proceedings of conference on,—London, November 1953 417
and electron diffraction photographs, Line interaction in 19, 270
determination of crystallinity in deformed natural rubber 321
diffraction rings, spotty, The determination of grain size from 257
diffraction work in the United States, Some thoughts on the future of natural science, with illustrations from the growth of 199
line breadths, The effect of intensifying screens on 454
scattering, Diffuse, and the physical properties of crystals 231

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

AUTHOR INDEX

Adkins, J. E., and A. N. Gent, Load-deflexion relations of rubber bush mountings 354
Allen, J. E., and J. D. Craggs, High current spark channels 446
Anderson, W., and E. H. Belcher, A standard light source of very low intensity based on the Čerenkov effect 53
Archard, G. D., Line interaction in X-ray and electron diffraction photographs 19, 270
Archard, G. D., Potential distribution of symmetrical cylindrical electron lenses 179
Archard, G. D., Requirements contributing to the design of devices used in correcting electron lenses 294
Archard, G. D., Some properties of symmetrical slit ("cylindrical") electron lenses 395
Archard, G. D., The effect of the finite light sources of measuring instruments on the determination of electron diffraction ring radii 69
Aris, R., *see under* Bosanquet C. H.
Arndt, U. W., and J. B. Nelson, Summarized proceedings of a conference on X-ray analysis—London, November 1953 417
Auld, J. H., *see under* Garrod, R. I.

Baker, B. O., Automatic electron trajectory plotting using the electrolytic tank analogue 191
Baker, S. C., Direct-reading metal spectroscopy with a d.c. arc 215
Bailey, R., *see under* Liebmann, G.
Barber, C. R., R. Handley, and E. F. G. Herington, The preparation and use of cells for the realization of the triple point of water 41
Belcher, E. H., *see under* Anderson, W.
Bertaud, S., *see under* M. C. Probine
Bett, K. E., K. E. Weale, and D. M. Newitt., The critical evaluation of compression data for liquids and a revision of the isotherms of mercury 243
Bigg, P. H., and F. H. Burch, Further tests on the stability of analytical weights in chemical laboratories 382
Blake, L. R., The determination of the texture of sheet from torque curves 99
Booker, G. R., A method of examining selected areas of surfaces using replicas and the electron microscope 349
Bosanquet, C. H., and R. Aris, On the application of Ångström's method of measuring thermal conductivity 252
Bradley, D. E., An evaporated carbon replica technique for use with the electron microscope and its application to the study of photographic grains 96
Bradley, D. E., Evaporated carbon films for use in electron microscopy 65
Brown, A. F. C., Triaxial stresses caused by notches 280
Bull, C. S., The noise of valves 127, 270
Burch, F. H., *see under* Bigg, P. H.
Burns, J. A., *see under* Emeleus, K. J.
Burrell, C. M., A simple bridge method of measurement of the time constant of exponential delays 196

Challice, C. E., Summarized proceedings of a conference on electron microscopy—London, November 1953 165
Chaundy, C. J. F., The anode fall in a glow discharge 255
Cochran, W., Summarized proceedings of a conference on X-ray analysis—Cambridge, April 1954 345
Cohen, H., *see under* Martin, B. W.
Collins, R. D., and H. Lubanska, The depression of liquid surfaces by gas jets 22
Cook, A. H., The calibration of circular scales and precision polygons 367
Craggs, J. D., *see under* Allen, J. E.
Craggs, J. D., *see under* McCormick, N. R.
Crangle, J., A magnetic method for the determination of gamma-loops in binary iron alloys and its application to the iron-silicon system 151
Crease, J., and M. J. Tucker, A note on an improved planimetric method of harmonic analysis 143
Curtis, G. C., The correction of an error in the determination of high resistances by capacitor discharge 106, 454

Dadson, R. S., D. W. Robinson, and R. G. P. Greig, The mechanical impedance of the human mastoid process 435
Daniel, V., The thermodynamic approach to solid structure 305
Davidson, P. M., The growth of the liquid bridge in an electrical contact 189
Davy, N., *see under* Langton, N. H.
Davy, N., *see under* Peake, H. J.
Davy, N., *see under* Snowdon, W.
Dawes, J. G., Densitometric evaluation of coal-dust stains on filter paper 221
Daws, L. F., The emissivity of a groove 182
De Boer, F., and W. F. Niklas, Application of a tracer to cathode-gettering and gas-adsorption problems 341
Denton, E. P., *see under* Rawson, H.
Donaldson, R., Spectrophotometry of fluorescent pigments 210
Douce, J. L., *see under* West, J. C.
Douglas, R. W., *see under* Goldsmid, H. J.
Duncan, J. F., and D. T. Warren, High vacuum applications of polythene 66
Dyke, V. E., The firing of pyrophylite for minimum distortion 270

Edels, H., *see under* Gambling, W. A.
Elliott, A., W. E. Hanby, and B. R. Malcolm, The near infra-red absorption spectra of natural and synthetic fibres 377
Emeleus, K. G., and J. A. Burns, Sub-normal discharges 277
Ennos, A. E., The sources of electron-induced contamination in kinetic vacuum systems 27
Evans, I., The rolling resistance of a wheel with a solid rubber tyre 187

Fremlin, J. H., *see under* Walker, D.
Frisch, O. R., Atomic energy—how it all began 81
Fuller, A. T., Control systems with quasi-critical damping 174
Fürth, R., Statistical concepts in theoretical physics 273

Gambling, W. A., and H. Edels, The high-pressure glow discharge in air 36
Garlick, G. F. J., *see under* Wright, G. T.
Garrod, R. I., and J. H. Auld, The effect of intensifying screens on X-ray line breadths 454
Gates, J. W., The evaluation of interferograms by displacement and stereoscopic methods 133
Gent, A. N., *see under* Adkins, J. E.
Gibbs, D. F., A simple bridge method of measurement of the time constant of exponential decays 196
Goldsmid, H. J., and R. W. Douglas, The use of semi-conductors in thermoelectric refrigeration 386, 458
Greig, R. G. P., *see under* Dadson, R. S.

Halmshaw, R., The use and scope of Iridium 192 for the radiography of steel 238

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

Hanby, W. E., *see under* Elliott, A.
Handley, R., *see under* Barber, C. R.,
Hardy, D. R., and H. Wroe, The electrical breakdown of pre-stressed air at atmospheric pressure 335
Harrison, E. R., Approximate electrode shapes for a cylindrical electron beam 40
Harrison, W., A check on the standard observer data at 4358 Å 269
Herington, E. F. G., *see under* Barber, C. R.
Herold, E. W., New advances in the junction transistor 115
Hickman, M. J., *see under* Powell, R. W.
Hird, D., *see under* Lawson, D. I.,
Hirsch, P. B., The determination of grain size from spotty X-ray diffraction rings 257
Holland, L., The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in vacuum 411
Hunt, G. H., The flux linkage with a search coil produced by a coaxial uniformly magnetized prolate spheroid 260

Jain, S. C., *see under* Krishnan, K. S.

Karbowiak, A. E., The elliptic surface wave 328
Krishnan, K. S., and S. C. Jain, Determination of thermal conductivities at high temperatures 426

Langridge, A., and M. P. Lord, The measurement of spectral distributions of luminescent materials above room temperature 196
Langton, N. H., and N. Davy, The two-dimensional electric field of a curved-sided wall or groove on an infinite plane 405
Langton, N. H., and N. Davy, The two-dimensional magnetic or electric field of a single isolated pole-piece 431
Lawson, D. I., and D. Hird, A photometric method of determining configuration factors 72
Liebmann, G., and R. Bailey, An improved experimental iteration method for use with resistance-network analogues 32
Liebmann, G., Note on the resistance-network analogue solution of field problems of spherical symmetry 412
Liebmann, G., Resistance-network analogues with unequal meshes or subdivided meshes 362
Link, W. T., *see under* Walker, D.
Litting, C. N. W., A method of investigating the secondary emissive properties of insulators 289
Llewellyn Jones, F., Ionization processes in the electrical breakdown of gases 49
Lord, M. P., *see under* Langridge, A.
Lubanska, H., *see under* Collins, R. D.

McCleery, D. K., The correction of an error in the determination of high resistances by capacitor discharges 454
McCormick, N. R., and J. D. Craggs, Some measurements of the relative dielectric strength of gases 171
McCree, K. J., The solarization of polymethylmethacrylate 454
MacEwan, D. M. C., Line interaction in X-ray and electron diffraction photographs 270
Mair, W. A., Supersonic gas flow 1
Malcolm, B. R., *see under* Elliott, A.
Mandel, L., Accuracy limitations of beta-ray thickness measurement 58
Mandel, L., The beta-ray absorption spectrum of $^{147}Pm_{61}$ and its application to thickness measurement 287
Martin, B. W., and H. Cohen, Heat transfer by free convection in an open thermosyphon tube 91
Morris, A. R., Some observations on the time edge effect in Catalin 800 104

Nelson, J. B., *see under* U. W. Arndt
Newitt, D. M., *see under* Bett, K. E.
Niklas, W. F., *see under* De Boer, F.
Nyburg, S. C., X-ray determination of crystallinity in deformed natural rubber 321

Oatley, C. W., The experimental determination of the speed of a vacuum pump and of components of a vacuum system 358

Parkinson, D. H., and J. E. Quarrington, The specific heat of polymerized Araldite and Wood's metal between 1.5 and 20° K 219
Peake, H. J., and N. Davy, The capacity and field of a split cylindrical condenser, using the method of inversion 316
Peake, H. J., and N. Davy, The capacity and field of a split cylindrical condenser when conductors differ in length 371
Pearce, C. A. R., The mechanism of the resolution of water-in-oil emulsions by electric treatment 136
Pengelly, A. E., and D. A. Wright, The emission from hot cathodes in gas discharges 391
Powell, R. W., and M. J. Hickman, Methods for studying the thermal resistances of sprayed and electrodeposited metal coatings 312
Price, P. H., and M. R. Slack, The effect of latent heat on numerical solutions of the heat flow equation 285
Probine, M. C., and S. Bertaud, A note on the selection of observers for primary standard optical pyrometry 227
Pyatt, E. C., Some consideration of the errors of brightness and two-colour types of spectral radiation pyrometer 264

Quarrington, J. E., *see under* D. H. Parkinson

Rapier, A. C., A theoretical investigation of the temperature distribution in the metal cutting process 400
Rawson, H., and E. P. Denton, The glass sealing properties of titanium and zirconium 352
Roberts, M. H., A potentiostat for corrosion study 351
Robinson, D. W., *see under* Dadson, R. S.
Rossman, M. G., and J. Yarwood, The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in a vacuum 7

Sack, R. A., Transverse oscillations in travelling strings 224
Slack, M. R., *see under* Price, P.H.
Smith, D. A., An improved technique for the microelectrophoresis of oil drops 325
Smith, E. V., Reflectance measurements on siliceous dusts and powders 340
Snowdon, W., and N. Davy, An application of conformal transformation to the investigation of the magnetic field between galvanometer pole-pieces 146
Stephens, K. G., *see under* Walker, D.
Stops, D. W., The molecular flow of gases 269

Tucker, M. J., *see under* Crease, J.

Vlam, C. C., The shape of the emission bands of luminescent solids 443

Walker, D., J. H. Fremlin, W. T. Link, and K. G. Stephens, The acceleration of heavy ions in a fixed-frequency cyclotron 157
Ward, A. G., The physical properties of gelatin solutions and gels 85

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

Warren, D. T., *see under* Duncan, J. F.
Weale, K. E., *see under* Bett, K. E.
West, J. C., and J. L. Douce, The frequency response of a certain class of non-linear feedback systems 204
Woods, J., and D. A. Wright, The thermionic emission from thin films of barium and strontium oxide 74
Wooster, W. A., Diffuse X-ray scattering and the physical properties of crystal 231
Wright, D. A., Some effects of slow electron bombardment in thermionic valves 108
Wright, D. A., *see under* Pengelly, A. E.
Wright, D. A. *see under* Woods, J.
Wright, G. T., and G. F. J. Garlick, Characteristics of radioluminescence in crystals 13
Wroe, H., *see under* Hardy, D. R.
Wyckoff, R. W. G., Some thoughts on the future of natural science, with illustrations from the growth of X-ray diffraction work in the United States 199

Yarwood, J., The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in vacuum 411
Yarwood, J., *see under* Rossman, M. G.

For *The Physics of Particle Size Analysis* and cognate subjects *see also* separate index to the supplement

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

SPECIAL ARTICLE

Supersonic gas flow*

By Prof. W. A. Mair, M.A., A.F.R.Ae.S., Engineering Laboratory, Cambridge

An introduction to supersonic gas flow is given, with emphasis on the features that distinguish supersonic from subsonic flow. Experimental methods for the study of supersonic flow are briefly considered and some simple flows are described.

In supersonic gas flow the velocity of the gas stream, relative to the solid boundaries, is greater than the velocity of sound in the gas. The object of this article is to describe some of the more important phenomena in supersonic gas flow, and to explain the principal differences between supersonic and subsonic flow. There have been extensive developments in this subject in recent years, but in a short article of this kind only an elementary treatment is possible.

The early work on supersonic flow was concerned mainly with problems of external ballistics, and comparatively little was done until developments in aeronautics stimulated further interest about ten years ago. Since then the subject has been much studied, both for its intrinsic physical interest and because of its aeronautical applications.

THE MACH NUMBER AND ITS SIGNIFICANCE

Considering fields of flow with geometrically similar boundaries, the theory of dimensions shows that any *dimensionless* property of the flow, e.g. a velocity ratio or a pressure ratio, can only depend on a set of *dimensionless* parameters that are functions of the independent variables. Moreover, these dimensionless parameters are *criteria of similarity*, i.e. if they have the same values in two fields of flow the fields are dynamically similar. It will be shown later that in the flow of a gas at speeds that are small compared with the velocity of sound the compressibility of the gas has no appreciable effect. For this case and for the flow of liquids (having negligible compressibility) the independent variables are simply the density ρ, the viscosity μ, the velocity V at some standard reference point in the field, and a length L defining the scale of the system. The only criterion of similarity is then the Reynolds number $R = \rho VL/\mu$. (It is assumed here that heat conduction and gravitational forces have no appreciable effect, and that the molecular mean free path is very small compared with the length L.)

In cases where the compressibility of the gas is important, the bulk modulus K and the ratio of specific heats γ must be included as independent variables. (The bulk modulus K is defined as $\rho(dp/d\rho)$ and for reversible adiabatic changes is equal to ρa^2, where a is the velocity of sound in the gas.) There are now two further criteria of similarity, in addition to the Reynolds number. The first of these is simply the ratio of specific heats γ; the second is the Mach number M, defined as

$$M = (\rho V^2/K)^{\frac{1}{2}} = V/a$$

where a is the velocity of sound.

Each of these criteria of similarity, Reynolds number, Mach number, and ratio of specific heats, is important in determining the flow of a gas at speeds where the compressibility cannot be neglected. The Mach number is particularly important in high-speed gas flow because, as is shown below, it determines the magnitude of the density changes associated with the pressure variations in the flow field. Moreover, there are important changes in the properties of the flow as the Mach number passes through 1, i.e. as the velocity of flow changes from subsonic to supersonic. The Reynolds number determines the relative magnitude of the viscous and inertia forces in the gas; many of the effects of varying Reynolds number are qualitatively the same in supersonic flow as at low speeds.

The relation between the Mach number and the density changes due to compressibility will now be considered. If V is the velocity of flow and ρ is the density of the gas, the changes of pressure in the flow (say Δp) are of order ρV^2. (This may be seen, for example, by considering Bernoulli's equation for the flow of an incompressible fluid, $p + \frac{1}{2}\rho V^2 =$ constant.) The associated density change ($\Delta\rho$) is of order $\rho V^2/(dp/d\rho)$. But for reversible adiabatic changes $dp/d\rho = a^2$, where a is the velocity of sound, and $dp/d\rho$ is of the same order for any type of change. Thus

$$\frac{\Delta\rho}{\rho} \text{ is of order } \frac{V^2}{a^2} = M^2$$

In particular, if $M^2 \ll 1$, $\Delta\rho \ll \rho$ and the behaviour of the gas is very nearly the same as that of an incompressible fluid. This result, which can also be obtained formally from the equations of motion, is of great importance because it means that, if the Mach number is small, compressibility has no appreciable effect. The bulk modulus K and the ratio of specific heats γ can then be omitted from the list of independent variables, and M and γ no longer appear as criteria of similarity. Thus experiments to investigate a flow at a low Mach number can be made at *any* small Mach number (and in a gas having any value of γ). The only criterion of similarity is then the Reynolds number.

Using a simple result from the kinetic theory of gases, the Mach number can be related to the molecular velocity. The pressure of a gas is given by

$$p = \tfrac{1}{3}\rho C^2$$

where C is the r.m.s. velocity of the molecules. The velocity of sound then becomes

$$a = (\gamma p/\rho)^{\frac{1}{2}} = C(\gamma/3)^{\frac{1}{2}}$$

and the Mach number is

$$M = V/a = (V/C)(3/\gamma)^{\frac{1}{2}}$$

Thus the Mach number is directly proportional to the ratio of the stream velocity to the molecular velocity. An alternative statement of this result is that the square of the Mach

* Based on a lecture given to the North Eastern and Manchester branches of The Institute of Physics on 18 and 20 March, 1953.

number is proportional to the ratio of the kinetic energies associated with the gas flow and with the thermal agitation of the molecules.

SHOCK WAVES

One of the important differences between supersonic and subsonic flow is that in a supersonic stream it is possible for stationary shock waves to be formed, whereas, in an entirely subsonic stream this is not possible. A shock wave is a thin layer (in motion relative to the gas) in which the velocity of flow and the state of the gas change suddenly. Neglecting the effects of viscosity and thermal conductivity, it can easily be shown that in a supersonic stream the energy, momentum, and continuity equations can be satisfied across a surface of discontinuity, i.e. a layer of zero thickness. The equations show that there are apparently two possibilities. Either the velocity component normal to the shock wave may decrease from a supersonic to a subsonic value as the gas passes through the shock wave, with a corresponding increase of pressure, density, and temperature, or alternatively all the changes may be in the reverse direction. Further investigation shows, however, that only the former alternative is possible, as the latter would involve a loss of entropy and this is known to be impossible in adiabatic conditions. In fact, there is a gain of entropy at a shock wave; the change is irreversible and there is a dissipation of energy.

For a more detailed investigation of the flow in a shock wave it is necessary to consider the effects of viscosity and thermal conductivity. It is then clear that the shock wave cannot have zero thickness and by making some assumptions it can be shown that the thickness t is given roughly by

$$t \simeq 10\nu/Sa$$

where ν is the kinematic viscosity of the gas, a is the velocity of sound and S is the strength of the shock wave (defined as $\Delta\rho/\rho$, where ρ is the density). The assumptions on which the above equation is based are valid only if the shock wave is fairly weak (say $S < 1$). For stronger shock waves, more detailed investigations show that the thickness is not much greater than the molecular mean free path. It should be noted that even for a weak shock wave the thickness is very small. In a typical case with $S = 0{\cdot}1$, for a shock wave formed in an air stream at a Mach number of 2, the calculated thickness is about 0·01 mm.

As the gas passes through a shock wave, the velocity component normal to the shock wave is reduced. If the shock wave is not normal to the stream there is also a velocity component in the plane of the shock wave, and this component does not change. Thus for an oblique shock wave the inclination of the resultant velocity to the normal to the shock wave is greater on the downstream side of the shock wave than on the upstream side. This change of direction of flow at an oblique shock wave is of great importance and is discussed later in this article.

Since the velocity component normal to a shock wave must always be supersonic on the upstream side, it is clear that in an entirely subsonic stream no stationary shock waves can be formed. Moreover, the velocity of propagation of a moving shock wave, relative to the gas in front of it, must always be greater than the velocity of sound. The difference between the velocity of propagation and the velocity of sound increases with the strength of the shock wave. For very weak shock waves the difference is zero, i.e. a very weak shock wave is propagated into the undisturbed gas at the velocity of sound.

THE MACH CONE AND MACH ANGLE

A *slender* body in a gas stream produces only weak disturbances, and if the body is sufficiently slender it may be assumed that there are no strong shock waves and that all disturbances are propagated at the velocity of sound.

Now consider such a slender body at a point P in a uniform stream of gas of velocity V (Fig. 1). At successive unit time intervals the wavefronts will be spheres of radii a, $2a$, $3a$, etc., where a is the velocity of sound. Since all the disturbances are convected downstream with velocity V, the centres of the successive spheres are displaced from the position of the body by distances V, $2V$, $3V$, etc., to the points A_1, A_2, A_3, etc. If $V < a$ the flow is subsonic, as in (*a*), while if $V > a$ the flow is supersonic, as in (*b*). The diagrams show that in subsonic flow the disturbances ultimately reach all parts of the field, while in supersonic flow only the region within the cone BPC is affected by the body. The semi-apex angle μ of this cone is called the Mach angle. Fig. 1(*b*) shows that

$$\sin \mu = a/V = 1/M$$

where M is the Mach number.

In supersonic flow the disturbances reinforce each other

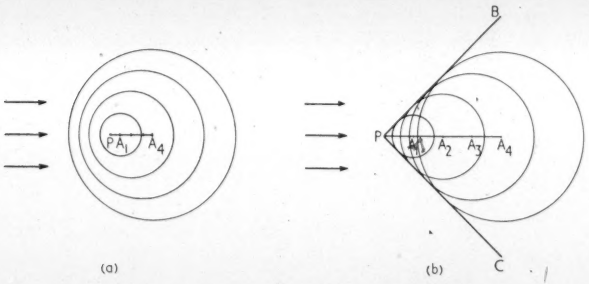


Fig. 1. Spherical wavefronts due to a slender body

(*a*) subsonic speed ($V < a$). (*b*) supersonic speed ($V > a$).

at the surface of the cone *BPC*, and a conical shock wave is usually formed. If the body is not slender, as assumed above, the shock wave is a strong one. The velocity component normal to the shock wave, on the upstream side, is then considerably greater than the velocity of sound, and the semi-apex angle of the conical shock wave is greater than the Mach angle. For a very slender body, however, the shock wave is weak and is inclined to the stream at the Mach angle.

This difference between subsonic and supersonic flow, illustrated in Fig. 1, is of great importance. For any point *P* in a supersonic stream there is a *downstream Mach cone* as shown at *BPC* in Fig. 1(*b*). There is also an *upstream Mach cone*, having the same apex angle, formed by producing the lines *BP* and *CP* on the upstream side of *P*. The following two rules can then be formulated:

(i) Any source of small disturbances (e.g. a slender body) in a supersonic stream can only affect the part of the field that is within the downstream Mach cone of the source. The source cannot have any effect on parts of the field outside this cone.

(ii) Any point in a supersonic stream can only be affected by disturbances originating within the upstream Mach cone of the point.

In a subsonic stream there are no such limitations; a disturbance at any point affects the whole field of flow to some extent.

It should be emphasized that the above conclusions are only true if the disturbances are all *small*, i.e. if there are no strong shock waves.

Fig. 1(*b*) shows that in a supersonic stream any *weak* disturbance (either compressive or expansive) is propagated along surfaces that are inclined to the stream at the Mach angle μ. This result has been obtained by considering a uniform stream, but in fact it is correct for any steady flow. In the general case, when the Mach number varies between one part of the field and another, the Mach angle must also vary. In the case shown in Fig. 1(*b*) the flow is axially symmetrical and the surface along which the disturbance is propagated is conical. In two-dimensional flow any disturbance is propagated along two planes intersecting at the source of the disturbance; the lines where these planes cut the plane of reference are called *Mach lines*. In axially symmetrical flow the Mach lines are the lines such as *PB* and *PC* in Fig. 1(*b*). Since for any position of a disturbance there are two Mach lines, the whole field is covered by two families of such lines. Consideration of Fig. 1(*b*) shows that the velocity component perpendicular to any Mach line must be equal to the local velocity of sound, and any very *weak* shock wave must lie along a Mach line.

If the surface of one of the boundaries of a two-dimensional supersonic stream is roughened, the disturbances produced by the roughness are propagated along the Mach lines and can be observed or photographed by one of the optical methods described later in this article. Thus the inclination of the Mach lines to the stream can be determined experimentally; this is the basis of a possible method of measuring Mach number in a uniform stream. Alternatively, instead of roughening the boundary, a slender pointed probe can be inserted into the stream. If the probe is *slender*, it produces a very weak conical shock wave whose semi-apex angle is equal to the Mach angle μ. (If the probe is not slender it produces a shock wave of appreciable strength, and the semi-apex angle of the cone is greater than μ.)

An important feature of the two families of Mach lines (in two-dimensional or axially symmetrical flow) is that they provide a network along which it is possible to integrate the equations of motion for an inviscid gas. This is the basis of the method of characteristics, a step-by-step numerical method of calculating inviscid flow with supersonic velocities throughout the field.

It has sometimes been stated that the shock waves formed by a body in a supersonic stream are analogous to the waves formed by a ship in deep water. This is not correct; in the case of a ship the wave velocity depends on the wavelength, which in turn depends on the velocity of the ship. The full analysis of the motion is much more complicated than in the case of supersonic gas flow, but it is found that the bow wave of a ship is inclined to the direction of motion at an angle that is nearly independent of the speed of the ship. In supersonic gas flow the wave velocity is constant (for small amplitudes), so that the wave inclination depends on the velocity of flow in the manner already explained.

Although the waves that are formed in *deep* water (e.g. by a ship) have no analogy in gas flow, there is an analogy between wave motion in *shallow* water and two-dimensional gas flow. The essential condition is that the depth of the water should be very small compared with the wavelength. If "ripples" having very small wavelengths are disregarded, the effects of surface tension may be neglected and the wave velocity is then $(gh)^{\frac{1}{2}}$, where h is the depth. Thus in these conditions the wave velocity is independent of wavelength and, if viscosity is negelcted, it can be shown that the motion of shallow water is analogous to the two-dimensional flow of a gas with $\gamma = 2$. With the limitations mentioned above, the velocity of surface waves on the water corresponds to the velocity of sound in the gas, and depth ratios in the water correspond to density ratios in the gas. The phenomenon in water that is analogous to a shock wave in a gas is the "hydraulic jump," but here the analogy is not an exact one unless the jump or shock wave is very weak.

The hydraulic analogy can be usefully employed for demonstrating some of the phenomena of supersonic gas flow, but it has little value for quantitative investigations because of the effects of surface tension and viscosity.

It has now been shown that there are some important differences between subsonic and supersonic gas flow. The following account of flow in ducts illustrates some further differences.

FLOW IN DUCTS

It is of interest to consider the adiabatic flow of a gas in a duct of varying area, in the absence of shock waves and neglecting the effects of viscosity. With these assumptions there can be no change of entropy, so that the flow is described as isentropic. Unless the scale is very small or the pressure very low, the Reynolds number of any supersonic gas flow is large. Because of this, the effects of viscosity are usually small except in long ducts or in shock waves, boundary layers, or wakes. Thus with these limitations the results of calculations with the assumptions given above may be applied to any stream tube in a real gas.

The theory given below is one-dimensional, i.e. it is assumed that conditions are uniform across any section of the duct and that only the component of velocity parallel to the axis need be considered. The continuity equation, expressing the condition for constant mass flow, is then

$$\rho V A = \text{constant} \qquad (1)$$

where ρ is the density and V the velocity at any section where the area is A.

If $M^2 \ll 1$ the gas may be regarded as incompressible, i.e. the density ρ is constant. Equation (1) then shows that the velocity is inversely proportional to the area of cross-section.

If M^2 is not small compared with 1, the density ρ must be regarded as a variable. Then equation (1) may be differentiated to give

$$\frac{d\rho}{\rho} + \frac{dV}{V} + \frac{dA}{A} = 0 \qquad (1a)$$

The one-dimensional momentum equation (sometimes known as Euler's equation) is

$$\frac{dp}{\rho} + VdV = 0 \qquad (2)$$

where p is the pressure.

Since the flow is assumed to be adiabatic and isentropic, the pressure-density relationship is the same as in a sound wave, so that

$$\frac{dp}{d\rho} = a^2 \qquad (3)$$

where a is the velocity of sound.

Equation (2) then becomes

$$a^2\frac{d\rho}{\rho} + VdV = 0 \qquad (4)$$

Substituting for $d\rho/\rho$ from equation (1a) then gives

$$\frac{dA}{A} = \frac{dV}{V}\left[\left(\frac{V}{a}\right)^2 - 1\right] = \frac{dV}{V}(M^2 - 1) \qquad (5)$$

At first sight, some of the conclusions from this equation are surprising. If the flow is subsonic dA/dV is negative, as for incompressible flow, so that the gas accelerates in a converging duct, and *vice versa*. If the flow is supersonic dA/dV is *positive*, and the gas *accelerates* in a *diverging* duct. The physical explanation of this is that if the flow is supersonic the reduction of density associated with an increase of velocity is so large that the required cross-sectional area actually increases with velocity.

Equation (5) also shows that $dA/dV = 0$ when $M = 1$. This means that, with these assumptions, sonic velocity can only be reached either in a parallel duct or at the section of minimum area (the "throat") in a convergent-divergent nozzle. Fig. 2 shows the calculated pressure, temperature, and area ratios in a convergent-divergent nozzle, as functions of Mach number. In this diagram, suffix 0 refers to the conditions in the supply reservoir, where the velocity is effectively zero, and suffix S refers to the sonic "throat." The conditions shown in Fig. 2 can only be realized if the overall pressure ratio of the nozzle is large enough; the flow is then subsonic in the converging part, sonic at the "throat," and supersonic in the diverging part.

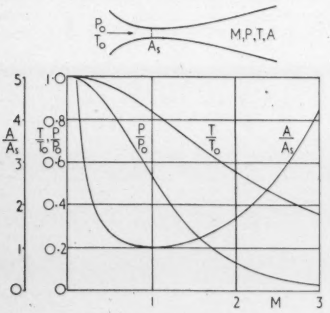


Fig. 2. Isentropic flow in a nozzle

The area changes discussed above illustrate one of the important differences between subsonic and supersonic flow. Because of this difference, the streamline pattern in the flow past a given body undergoes large changes as the velocity of flow increases from a subsonic to a supersonic value.

The flow in a convergent-divergent nozzle with varying back pressure will now be considered. It will be assumed that the flow is adiabatic and isentropic, except where shock waves occur. In Fig. 3, the abscissa is the distance along the nozzle, measured from the inlet, P is the local pressure at any section, P_0 is the pressure in the supply reservoir, as in Fig. 2, and P_2 is the pressure outside the nozzle exit. If P_2 is not greater than the pressure at the point C, the pressure distribution along the nozzle is as shown by the full line ABC. There is then a continuous increase of velocity through the nozzle, as in Fig. 2. If P_2 is equal to the pressure at the point D, the pressure distribution is as shown by the full line ABD, and the flow in the nozzle is entirely subsonic. The velocity increases up to the velocity of sound in the convergent part of the nozzle, then decreases again in the divergent part. For higher values of P_2 the velocities are still subsonic throughout, but the maximum velocity (occurring at the throat) is less than the velocity of sound.

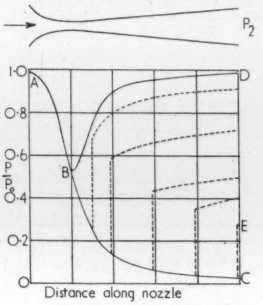


Fig. 3. Flow in a convergent-divergent nozzle with variable back pressure

If P_2 has a value corresponding to a point between C and D, one or more shock waves are formed in the divergent part of the nozzle or at the exit. The dotted curves in Fig. 3 show alternative pressure distributions for various positions of a normal shock wave. (Because of the boundary layer on the wall, the pressure at the wall actually rises more gradually at the shock wave than is shown by these curves.) The dotted curves show that flow with a normal shock wave is possible for values of P_2 corresponding to points between D and E. For values corresponding to points between E and C, a system of *oblique* shock waves is formed at the nozzle exit.

A supersonic wind tunnel consists essentially of a convergent-divergent nozzle, with a compressor or other means

of maintaining the required overall pressure ratio. It is necessary to shape the divergent portion of the nozzle correctly to give uniform flow. The correct shape can be calculated (for inviscid flow) by the method of characteristics already mentioned.

SOME EXAMPLES OF SUPERSONIC FLOW

It has already been explained that there is a change of stream direction at an oblique shock wave, the streamlines being inclined to the normal at a greater angle on the downstream side of the shock wave than on the upstream side. By considering this change of direction, it can be shown that in some conditions the supersonic flow of a gas past a two-dimensional wedge is particularly simple. Fig. 4(*a*) shows a wedge of apex angle 30° in a stream of Mach number 2. The shock waves *AB* and *AC* are attached to the apex of the wedge, and have a strength such that they deflect the stream

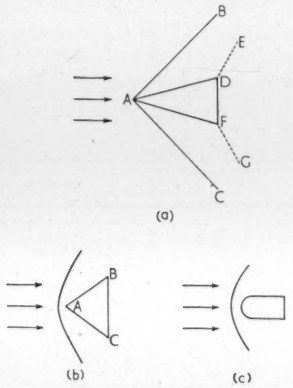


Fig. 4. Supersonic flow past two-dimensional bodies

exactly 15° (half the apex angle of the wedge). Behind these shock waves the pressure, velocity, and stream direction are all uniform as far as the Mach lines *DE* and *FG*, where the region affected by the back of the wedge begins. The wedge has no effect on the flow upstream of the shock waves *AB* and *AC*, but at these shock waves there is a sudden change of stream direction.

The flow is qualitatively similar to that shown in Fig. 4(*a*) for a wide range of wedge angles and Mach numbers. As the wedge angle is increased, however, a stage is reached at which there is no possible shock wave giving the required deflexion of the stream. (This also occurs if the Mach number is reduced, since there is then a corresponding reduction of the maximum possible stream deflexion at a shock wave.) Under these conditions the shock wave becomes detached from the apex of the wedge as in Fig. 4(*b*). Since the part of the shock wave immediately in front of the apex *A* is normal to the stream, the velocities near the apex must be subsonic, and in fact the velocity along the surface of the wedge remains subsonic as far as the shoulders *B* and *C*.

The flow past a cone is in some respects similar; there is a conical shock wave attached to the apex of the cone if the cone angle is not too large and the Mach number not too small. In this case, however, the stream lines are curved after passing through the shock wave, and the deflexion at the shock wave is *less* than the semi-apex angle of the cone.

If a blunt-nosed body is placed in a supersonic stream the shock wave is always detached from the nose, as in Fig. 4(*c*), because the required stream deflexion at an attached shock wave would be 90°, and this is never possible. Considering the equivalent case of a blunt-nosed body moving at supersonic speed through stationary air, the shock wave remains at a fixed distance in front of the body, and hence must advance into the undisturbed air at the same speed as the

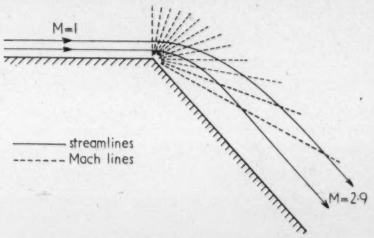


Fig. 5. Supersonic flow around a corner (deflexion = 48°)

body. The strength of the shock wave adjusts itself to satisfy this condition.

An interesting example of a supersonic flow for which there is a simple theoretical solution (neglecting viscosity) is the flow around a convex corner in two dimensions. In subsonic flow, inviscid theory predicts an infinite velocity at such a sharp corner, and in a real fluid the flow separates from the boundary at the corner. In supersonic flow the streamlines are as shown in Fig. 5; there is a continuous fall of pressure and increase of velocity, and no tendency for the flow to separate from the surface.

In Fig. 5 the streamlines and Mach lines are shown for a Mach number of 1 on the upstream side of the corner and 2·9 on the downstream side. The state of the gas and the

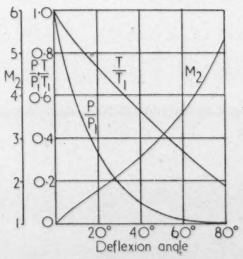


Fig. 6. Conditions in supersonic flow around a corner (initial Mach number 1)

velocity are constant along any Mach line. It is found from the theory that if the gas behaved as a perfect gas at all temperatures, a deflexion of 130° would increase the Mach

number from 1 to infinity. The velocity on the downstream side would still be finite, but the absolute temperature would be zero, giving zero velocity of sound.

With a Mach number of 1 on the upstream side, the angle of deflexion could have any value between 0 and 130° in this type of flow. In Fig. 6 the temperature and pressure ratios and the Mach number M_2 on the downstream side are plotted against the angle of deflexion. (Suffix 1 refers to the conditions on the upstream side, where $M = 1$.)

A flow of this kind is of course also possible with the upstream Mach number greater than 1. In this case the deflexion required to produce specified conditions is found from Fig. 6 by subtracting the deflexion angles corresponding to the upstream and downstream Mach numbers.

EXPERIMENTAL METHODS

It has been explained that any body placed in a supersonic gas stream produces shock waves. One consequence of this is that shock waves are formed when any measuring instrument is inserted into the stream, and these shock waves may affect the quantities to be measured. Thus, although the conventional flow-measuring instruments, such as Pitot and static tubes, can sometimes be used in supersonic flow, there is a need for methods of investigation that do not require the insertion of any body into the stream. Fortunately the changes of density, that are an essential feature of supersonic flow, are associated with changes of refractive index of the gas, and these changes can be observed by several optical methods.

The simplest of these methods is the simple shadowgraph (Fig. 7). A parallel beam of light is passed through the gas

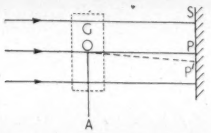


Fig. 7. Principle of shadowgraph

stream to be investigated G and illuminates a screen or photographic plate S. If, in a region O, there is a density gradient with a component in the direction OA, a ray of light that would reach the point P in the absence of the density gradient is deflected to the point P'. Thus there is a dark region on the screen at P and a light region at P'. It can easily be shown that with this arrangement the illumination on the screen depends on the rate of change of density gradient, and since this quantity becomes very large at a shock wave the system is excellent for showing the positions of shock waves in the flow.

Another method, more suitable for detecting small density gradients, is the so-called "schlieren" method suggested by Toepler. Fig. 8 shows a typical form of the apparatus, using two concave mirrors. (There are many other possible arrangements, but the principles of these are not essentially different.) In Fig. 8, light from the source S is focused by the lens L_1 on to the slit T. This slit is in the focal plane of

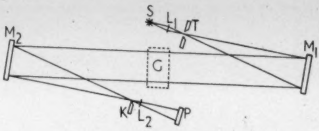


Fig. 8. Schlieren apparatus

the mirror M_1, so that a parallel beam of light passes through the gas stream G to be investigated. A knife edge K is placed in the focal plane of the mirror M_2, so that an image of the illuminated slit T is formed at K, part of the image of the slit being obstructed by the knife edge. If there is a density gradient in the gas stream having a component perpendicular to the knife edge K, the light beam will be refracted either towards the knife edge or away from it, and there will be a change of illumination on the screen at P. The convex lens L_2, in conjunction with the mirror M_2, gives a real image of an object in the gas stream on the screen P. Thus refraction of a ray of light in the gas stream does not affect the position of the point where the ray reaches the screen, but it does affect the intensity of illumination at that point.

Using the schlieren method, photographs can be taken to show the positions of the regions where there is an appreciable density gradient, but because of diffraction effects the method cannot be used quantitatively to determine the magnitudes of the density gradients. Quantitative measurements can be made, however, by using an interferometer; the variations of density in the gas flow can be calculated from the observed fringe shifts.

BIBLIOGRAPHY

LIEPMANN, H. W., and PUCKETT, A. E. *Introduction to Aerodynamics of a Compressible Fluid.* (New York: John Wiley and Sons, Inc., 1947).

DURAND, W. F. (Editor). *Aerodynamic Theory.* Division H, Vol. 3, *The Mechanics of Compressible Fluids.* By G. I. Taylor and J. W. Maccoll (Berlin: Springer, 1935).

FERRI, A. *Elements of Aerodynamics of Supersonic Flows.* (New York: The Macmillan Co., 1949).

SAVER, R. *Introduction to Theoretical Gas Dynamics.* English translation by F. K. Hill and R. A. Alpher (Ann Arbor: J. W. Edwards, 1947).

The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in a vacuum

By M. G. Rossmann, M.Sc., Grad.Inst.P., Natural Philosophy Department, The Royal Technical College, Glasgow, C.1, and J. Yarwood, M.Sc., F.Inst.P., Department of Mathematics and Physics, The Polytechnic, Regent Street, London, W.1

[*Paper first received* 18 *May, and in final form* 5 *August*, 1953]

The conditions are discussed which need to be fulfilled to obtain measurements on the evaporation of metals which liquefy when at elevated temperatures in a vacuum. A new type of crucible for such work is described. The impedance of a crucible to flow of metal vapour is considered theoretically. Experimental results for the evaporation rates of mercury, silver and gold are given.

In order to evaporate a metal in a vacuum where the metal is in the liquid phase at the temperature concerned it is necessary, to obtain measurements of the evaporation temperature and of the rate of loss of mass per second per unit area, to fulfil the following requirements:

(*a*) The crucible material itself must not evaporate significantly or give rise to loose particles.
(*b*) There must be no chemical reaction between the material of the crucible and the hot metal.
(*c*) The surface of the liquefied metal must be of defined geometry and must remain constant in shape.
(*d*) The liquid metal must not be subject to bubbling.
(*e*) The temperature of the metal must be uniform and susceptible to ready determination.
(*f*) A knowledge of the ratio of the mass of vapour which could leave the surface per second to the mass escaping from the crucible must be obtained.

It should be asserted at the outset that it is practically impossible to fulfil all these needs properly. It is for this reason that there are no data available giving directly determined rates of evaporation for liquid metals at elevated temperatures.

A type of crucible which has been used with some success in our experiments is a double-walled crucible of carbon of a hollow cylindrical shape with a closed base. Such crucibles were heated with 0·5 mm or 1·0 mm diameter tungsten wire spirals wound directly round the carbon (Fig. 1). The fact

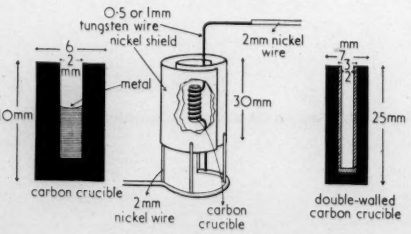


Fig. 1. Double-walled carbon crucible with heater and radiation shield

that the tungsten wire and the carbon were in direct electrical contact did not matter since the resistance of the spiral was considerably lower than the resistance *via* the carbon crucible. To conserve heat this arrangement was surrounded with a nickel heat shield of the form shown.

It is necessary to consider the virtues of this crucible in the light of requirements (*a*) to (*f*).

(*a*) The double-walled crucible consisting of one cylinder fitting closely inside the other was necessary to avoid loss of mass due to evaporation of the carbon immediately against the tungsten wire. The inner crucible, which could be removed and weighed with its contents, did not suffer this defect. In fact, blank experiments showed that the loss of weight of the inner crucible after heating in a vacuum for 1 hour at 1600° C was negligibly small. The crucibles were degassed by heating in a vacuum before use.

(*b*) As regards chemical action between the hot, molten metal and the crucible, there was no evidence of such action in experiments performed with gold and mercury, but evidence of the formation of silver carbide was occasionally observed when evaporating silver. This may partly account for the inconsistent results obtained for the evaporation rate of silver.

(*c*) In arriving at a figure for the rate of loss of mass of metal per cm^2 per sec, it is essential that a constant form of evaporating surface be maintained. Here the greatest difficulty occurred in experiments with aluminium. Molten aluminium wetted the surfaces of all the substances tried as crucibles, so that the evaporation area, and hence total rate of evaporation, increased enormously during the preliminary heating. Some experiments using specially prepared carbons formed from carbon black, and heated to over 2000° C during manufacture, did exhibit freedom from wetting with molten aluminium,* but the results were not entirely consistent. The difficulty of wetting was absent using silver, gold and mercury.

BOILING OF A LIQUID IN A VACUUM

(*d*) When considering measurements of the rate of evaporation of liquids in a vacuum it must be borne in mind that the reduced pressure may cause the liquid to boil at the temperature prevailing. Thus bubbles form within the liquid, and rising to the surface and bursting, modify greatly the surface geometry, thus giving completely false information concerning the rate of evaporation per unit area (*e*). In many circumstances, however, this effect can be neglected, as is shown below.

Let the incident pressure of the residual gas in the vacuum on the liquid surface be p cm of mercury (Fig. 2), and the vapour pressure of the liquid at the operating temperature be P cm of mercury. Suppose the density of the liquid is ρ g.cm^{-3} and the maximum depth of bubble formation is d cm.

A bubble cannot form if the vapour pressure P inside the

* The authors are indebted to Dr. M. Pirani for this suggestion.

potential bubble is smaller than the total incident pressure at the point in question. This incident pressure will be due to the gas pressure on the liquid surface plus the hydrostatic pressure due to the liquid above the point. It follows that d is given by an equation

$$P = p + \frac{\rho d}{13{\cdot}6} \tag{1}$$

In practice, p is very small (10^{-6} cm approx.) for a good vacuum, and may be neglected

Therefore

$$d = \frac{P \,.\, 13{\cdot}6}{\rho} \tag{2}$$

If the change of liquid surface form due to bubbling is to be negligible then d must be not greater than molecular dimensions. This would imply that P was of the order of

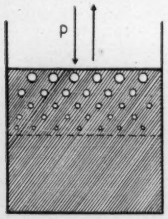


Fig. 2. Concerning bubble formation in the heated liquid

10^{-8} cm. For insignificantly small bubbles, however, values of P as great as 10^{-4} cm of mercury can be tolerated. At higher vapour pressures than this, however, the effect of bubbling becomes increasingly marked. This factor has been neglected by some writers who have calculated the rates of evaporation of liquid metals from their vapour pressures, e.g. Dushman.[1]

A liquid surface in a crucible will assume a curved form if wetting is absent. The effect of the change of vapour pressure with radius of curvature is negligible except for very small values of this radius, much less than those encountered in crucible experiments.

Concerning requirement (*e*) of the introductory paragraph, the use of carbon crucibles is particularly recommended because, under suitable conditions, they approach closely to black-body radiators, so enabling the temperature to be determined, with a calibrated optical pyrometer of the disappearing filament type, without the necessity of knowing the emissivity of the crucible material. Moreover, if the cylindrical hole in the crucible is not sufficiently deep for the assumption of black-body conditions to be justified, e.g. if the crucible is nearly filled with liquid metal, then the emissivity of the carbon can be found under the conditions prevailing in the experiment by sighting the pyrometer first on the empty cavity and, secondly, on the crucible mouth. The vacuum chamber used was provided with a flat glass window for observing the crucible with the pyrometer, a correction being applied for the heat transmission characteristics of the glass. For temperatures below 1000° C a thermocouple was preferable.

RELATION BETWEEN THE VAPOUR PRESSURE OF A SUBSTANCE AND ITS RATE OF EVAPORATION IN A VACUUM

It is usual to calculate the rate of evaporation of a substance from a knowledge of its vapour pressure in the following way:

Let η molecules leave unit plane area of the evaporating substance in unit time at a temperature of T° K. η will depend only on the temperature of the substance and the substance itself.

Let the effective loss of mass from unit area in unit time be μ, and the pressure of the substance's unsaturated vapour be p.

It can readily be shown, at low pressures, where the vapour may be regarded as obeying the perfect gas laws, that $p\sqrt{(M/2\pi RT)}$ g of the vapour hit unit area of the substance in unit time, where M is the gram-molecular weight of the substance and R is the molar gas constant.

Suppose that a fraction α of those molecules hitting the substance condense on it. α is now usually termed the condensation coefficient, but Knudsen calls it the accommodation coefficient. Then

$$\mu = m\eta - \alpha p\sqrt{(M/2\pi RT)} \tag{3}$$

where m is the mass of one molecule in grams.

The rate of evaporation will be zero when the vapour above the condensate is saturated, so that $p = P$, where P is the saturation vapour pressure of the substance.

Therefore

$$0 = m\eta - \alpha P\sqrt{(M/2\pi RT)}$$

Therefore

$$\mu = \alpha(P - p)\sqrt{(M/2\pi RT)} \tag{4}$$

This equation (4) reduces to

$$\mu = \alpha P\sqrt{(M/2\pi RT)} \tag{5}$$

if $P \gg p$. This condition prevails if the substance is evaporating in a high vacuum where the mean free path of the vapour molecules is so large that nearly all of them condense on the vessel walls before colliding with other molecules.

The usual method of finding μ, the rate of evaporation, is to determine P, by a method such as Knudsen's effusion method, and then substitute in equation (5) assuming that $\alpha = 1$. This last presumption is open to grave objections since α has not been determined for a large number of the substances for which the rates of evaporation, found from the vapour pressures, are given in standard works on vacuum physics.

THE IMPEDANCE OF THE CRUCIBLE TO THE FLOW OF METAL VAPOUR

With regard to requirement (*f*) of the introductory paragraph, a method of calculating the ratio of the amount of metal vapour which could leave the liquid surface to the metal vapour leaving the crucible orifice was partly based on the work of Buckley[2,3] who calculated the total heat radiation from the interior of an infinitely long cylinder of circular cross-section, and later from a cylinder of finite length. Metal atoms were considered in the present problem instead of heat radiation, as in Buckley's work, on the basis of the analogy between Knudsen's cosine distribution law for molecules and the cosine law for radiation, and also between the condensation coefficient in molecular phenomena and emissivity in radiation. The method used may well be valuable in other calculations concerning the movements of molecules along pipes and jets in vacuum physics.

The condensation coefficient is defined as

$$\frac{\text{the number of molecules which are condensed on hitting a surface}}{\text{total number of molecules hitting the surface}}$$

Emissivity is defined as

$$\frac{\text{the radiation which is absorbed by the surface}}{\text{total incident radiation on the surface}}$$

It is assumed that no, or negligibly few, molecules collide with each other within the crucible. In the experiments the mean free path was at least eight times the crucible diameter. Thus molecular streaming prevails. Under these conditions Knudsen[4] shows that the number of molecules reflected specularly is negligible.

Let the radius of the right circular cylindrical crucible be 1, the length of the crucible be l, and an annulus on the walls be defined by the vertical co-ordinate z, measured from the crucible base (Fig. 3).

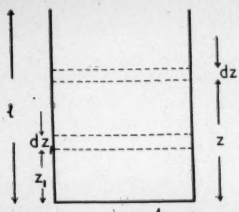


Fig. 3. Concerning the impedance of a cylindrical crucible to the flow of metal vapour

Let the condensation coefficient (or emissivity) of the base be α, and the condensation coefficient (or emissivity) of the walls be β. In the molecular case, $\beta = 0$ because no molecules condense on the walls, which are at the same temperature as the heated metal.

Let a molecules leave the base per unit area in unit time initially, and b molecules leave the walls per unit area in unit time initially.

In the radiation case $a = \alpha\delta T^4$, $b = \beta\delta T^4$, where δ is Stefan's constant. In the molecular case, $b = 0$.

Before proceeding any farther it is necessary to state two theorems.

Theorem 1[5]

The radiation from the inner walls of a cylinder between cross-sections A and B (Fig. 4) to the coaxial circular area C, is equal to that which C would receive from a circular disk

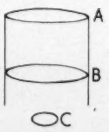


Fig. 4. Concerning the theorem that the radiation from the walls of a cylinder to a coaxial circular area C is equal to the radiation from disk A minus the radiation from disk B

at B less what it would receive from a circular disk at A, both disks having the same radiation rate (i.e. emit the same number of molecules in unit time).

Theorem 2

The radiation received by a circular disk from a diffusely radiating disk, parallel and coaxial with the receiving disk, is given by Walsh[6] as

$$\frac{\pi E}{2}\left\{(d^2 + r^2 + R^2) - [(d^2 + r^2 + R^2)^2 - 4r^2R^2]^{\frac{1}{2}}\right\}$$

where E is the total radiation emitted by the radiating disk per unit area (i.e. the number of molecules emitted in unit time per unit area), d is the distance apart of the disks, and r and R are their radii. In the present case $r = R = 1$, and $E = a$.

Consider two disks whose positions are given by co-ordinates z and z_1 (Fig. 3). Then the above expression reduces to

$$\frac{\pi a}{2}\left\{(z - z_1)^2 + 2 - (z - z_1)[(z - z_1)^2 + 4]^{\frac{1}{2}}\right\}$$

Put

$$F(z - z_1) = (z - z_1)^2 + 2 - (z - z_1)[(z - z_1)^2 + 4]^{\frac{1}{2}} \quad (6)$$

Thus $\frac{1}{2}\pi aF(z)$ molecules per unit time will hit a disk at height z, and $\frac{1}{2}\pi aF(z + dz)$ molecules per unit time will hit a disk at height $z + dz$, after emission from the base.

Therefore

$$\frac{\pi a}{2}[F(z) - F(z + dz)] = -\frac{\pi a}{2}\frac{d}{dz}F(z)dz$$

molecules will hit the annulus between z and $z + dz$.

Since the area of the annulus at height z and width dz is $2\pi \cdot 1 \cdot dz$, $-(a/4)(d/dz)F(z)$ molecules will hit the annulus at height z per unit area after being emitted from the base of the cylinder.

Hence, the total number of molecules which leave unit area in unit time from an annulus at z is now $\phi_0(z)$, where

$$\phi_0(z) = b - (1 - \beta)\frac{a}{4}\frac{d}{dz}F(z) \quad (7)$$

since b molecules are emitted from the walls of the cylinder per unit area on their own account, and $(1 - \beta)$ of all those which hit are reflected. The quantities b and β are preserved in this equation so that it can also be applied to the radiation case.

Let $\phi_1(z)$ be the number of molecules which are reflected from unit area, in unit time, from an annulus at z_1, which is being irradiated with molecules from the rest of the walls of the cylinder with a distribution $\phi_0(z)$.

Let us assume that $\phi_0(z)$ remains constant between z and $z + dz$.

Then the number of molecules received by a disk at z_1 from an annulus at z (see Fig. 3) is given by

$$\frac{\pi\phi_0(z)}{2}\left\{F(z \sim z_1) - F[(z + dz) \sim z_1]\right\}$$

according to Theorems 1 and 2.

$$= \frac{-\pi\phi_0(z)}{2}\frac{d}{dz}F(z \sim z_1)dz$$

Therefore, the number of molecules received by the annulus between z_1 and $z_1 + dz_1$ from an annulus at z is

$$\frac{-\pi\phi_0(z)}{2}\left\{\frac{d}{dz}F[z \sim (z_1 + dz_1)]dz - \frac{d}{dz}F(z \sim z_1)dz\right\}$$

$$= \frac{-\pi\phi_0(z)}{2}\frac{d}{dz_1}\left[\frac{d}{dz}F(z \sim z_1)dz\right]dz_1$$

Now $$\frac{d}{dz}F(z - z_1) = -\frac{d}{dz_1}F(z - z_1)$$

Therefore, the number of molecules received per unit area at z_1 from the annulus at z is given by

$$\frac{1}{2\pi dz_1} \cdot \frac{\pi\phi_0(z)}{2}\frac{d^2}{dz^2}F(z \sim z_1)dzdz_1$$

$$= \frac{\phi_0(z)}{4}\frac{d^2}{dz^2}F(z \sim z_1)dz$$

Writing $F(z - z_1)$, when $z > z_1$

and $F(z_1 - z)$, when $z_1 > z$

the total number of molecules reflected by an annulus at height z_1, after being emitted by the walls of the crucible with the distribution $\phi_0(z)$, is

$$\phi_1(z_1) = \frac{1-\beta}{4}\int_0^{z_1}\phi_0(z)\frac{d^2}{dz^2}F(z_1 - z)dz + \frac{1-\beta}{4}\int_{z_1}\phi_0(z)\frac{d^2}{dz^2}F(z - z_1)dz$$

Similar equations will be obtained for $\phi_2(z)$, $\phi_3(z)$, . . . Consequently, the final number of molecules which leave the annulus at z_1, due to inter-reflexion on the *walls* of the crucible, is

$$\Phi(z_1) = \phi_0(z_1) + \phi_1(z_1) + \phi_2(z_1) + \ldots$$

or

$$\Phi(z_1) = \phi_0(z_1) + \frac{1-\beta}{4}\int_0^{z_1}\Phi(z)\frac{d^2}{dz^2}F(z_1 - z)dz + \frac{1-\beta}{4}\int_{z_1}^{l}\Phi(z)\frac{d^2}{dz^2}F(z - z_1)dz \quad (8)$$

A similar integral equation was obtained and solved by a different method by Buckley. However, he was interested in determining $\Phi(z)$ for a *semi-infinite* cylinder. We are interested in the total number of molecules emitted from the mouth of a *finite* cylinder. To do this we will make use of equation (8), which is the reason why Buckley's work has been reproduced. Our calculations are shown below.

Now let ν_1 molecules hit the base per unit area after the first reflexion from the walls, and ν_2 molecules hit the base per unit area after the second reflexion from the walls, etc.

From Theorem 1 the number of molecules emitted from the walls between z and $z + dz$, due to the distribution $\phi_0(z)$ (assumed constant over dz), and which hit the base of the cylinder is

$$\frac{\pi\phi_0(z)}{2}[F(z) - F(z + dz)] = -\frac{\pi\phi_0(z)}{2}\frac{d}{dz}F(z)dz$$

Therefore, $$-\frac{1}{\pi \cdot l^2}\cdot\frac{\pi\phi_0(z)}{2}\cdot\frac{d}{dz}F(z)$$

molecules hit unit area of the base after being emitted from the walls of the crucible according to distribution $\phi_0(z)$, at height z.

Therefore, the total number of molecules which hit unit area of the base, after emission from the walls of the cylinder with distribution $\phi_0(z)$, is

$$\nu_1 = -\frac{1}{2}\int_0^{l}\phi_0(z)\frac{d}{dz}F(z)dz$$

Similarly $$\nu_2 = -\frac{1}{2}\int_0^{l}\phi_1(z)\frac{d}{dz}F(z)dz$$

and so on for ν_3, ν_4, . . . etc.

Therefore

$$\nu_1 + \nu_2 + \nu_3 + \ldots = n\ (\text{of } a) = -\frac{1}{2}\int_0^{l}\Phi(z)\frac{d}{dz}F(z)dz \quad (9)$$

Buckley[2] shows that, by the method of Prony,

$F(z) = 2e^{-z}$ to the first approximation;

$= 2{\cdot}42e^{-1{\cdot}14z} - 0{\cdot}42e^{-3{\cdot}4z}$ to the second approximation.

The first approximation is very good for $0 \leqslant z \leqslant 4$. For the purpose of this analysis, the first approximation only will be used.

Hence, from equation (8)

$$\Phi(z_1) = \phi_0(z_1) + \frac{1-\beta}{2}e^{-z_1}\int_0^{z_1}\Phi(z)e^{z}dz + \frac{1-\beta}{2}e^{z_1}\int_{z_1}^{l}\Phi(z)e^{-z}dz \quad (10)$$

where, from equation (7),

$$\phi_0(z) = b + (1-\beta)\frac{a}{2}e^{-z} \quad (11)$$

Consequently, from equation (9),

$$\nu_1 + \nu_2 + \nu_3 + \ldots = na = \int_0^{l}\Phi(z)e^{-z}dz \quad (12)$$

Let $$\Phi_1(z) = \int\Phi(z)e^{z}dz$$

and $$\Phi_2(z) = \int\Phi(z)e^{-z}dz$$

Then equation (10) can be written as

$$\Phi(z_1) = \phi_0(z_1) + \frac{1-\beta}{2}e^{-z_1}[\Phi_1(z_1) - \Phi_1(0)] + \frac{1-\beta}{2}e^{+z_1}[\Phi_2(l) - \Phi_2(z_1)]$$

On differentiating twice with respect to z_1, we obtain

$$\frac{d^2\Phi(z_1)}{dz_1^2} = \frac{d^2\phi_0(z_1)}{dz_1^2} - (1-\beta)\Phi(z_1) + \left\{\frac{1-\beta}{2}e^{-z_1}[\Phi_1(z_1) - \Phi_1(0)] + \frac{1-\beta}{2}e^{+z_1}[\Phi_2(l) - \Phi_2(z_1)]\right\}$$

But by differentiating equation (11) twice with respect to z we find that (on putting $z = z_1$)

$$\frac{d^2\phi_0(z_1)}{dz_1^2} = \phi_0(z_1) - b$$

Thus we have

$$\frac{d^2\Phi(z_1)}{dz_1^2} = \phi_0(z_1) - b - (1-\beta)\Phi(z_1) + \Phi(z_1) - \phi_0(z_1)$$

or $$\frac{d^2\Phi(z)}{dz^2} = \beta\Phi(z) - b \quad (13)$$

In the molecular problem $\beta = 0$ and $b = 0$, so equation (13) becomes

$$\frac{d^2\Phi(z)}{dz^2} = 0$$

Therefore $$\Phi(z) = A + Bz \quad (14)$$

where A and B are constants.

The integral equation (10) will now be

$$A + Bz_1 = \frac{a}{2}e^{-z_1} + \frac{e^{-z_1}}{2}\int_0^{z_1}(A + Bz)e^z dz + \frac{e^{z_1}}{2}\int_{z_1}^{l}(A + Bz)e^{-z}dz$$

Therefore

$$A + Bz_1 = A + Bz_1 + \frac{e^{-z_1}}{2}(a - A + B) - \frac{e^{-z_1-l}}{2}(A + Bl + B)$$

Thus A and B are given by the solution of two simultaneous equations

$$A - B = a$$

and

$$A + B(l + 1) = 0$$

Therefore $$A = \frac{a(l+1)}{l+2} \text{ and } B = -\frac{a}{l+2}$$

Hence $$\Phi(z) = \frac{a(l+1)}{l+2} - \frac{a}{l+2}z$$

Substituting this expression for $\Phi(z)$ in equations (12), and integrating between the limits, it will be found that

$$n = l/(l + 2) \tag{15}$$

Since n of a molecules hit the base per unit area, αna molecules condense on the base per unit area, and $(1 - \alpha)na$ molecules are reflected from the base per unit area. Therefore, n of $(1 - \alpha)na$ molecules hit the base again after inter-reflexion with the cylinder's walls. Of these, $\alpha n(1 - \alpha)na$ are condensed and $(1 - \alpha)n(1 - \alpha)na$ are reflected.

Therefore the total number of molecules which are condensed on the base per unit area is

$$\alpha na + \alpha n(1-\alpha)na + \alpha n(1-\alpha)n(1-\alpha)na + \ldots$$
$$= \alpha n[1 + n(1-\alpha) + (1-\alpha)^2n^2 + \ldots]$$
$$= \frac{\alpha na}{1-(1-\alpha)n} \text{ (summing the geometrical progression)}$$

i.e. the rate of evaporation, μ, is given by

$$\mu = a - \frac{\alpha na}{1-(1-\alpha)n}$$

or

$$\mu = \frac{a(1-n)}{1-(1-\alpha)n} \tag{16}$$

It should be noted that this geometrical progression is convergent only if $(1 - \alpha)n < 1$. This is true since n is always less than 1, because it is impossible that the number of molecules which return to the base of the crucible after inter-reflexions between the walls is greater than the number which leave the base per second. Also $0 < \alpha \leqslant 1$, and so $0 \leqslant (1 - \alpha) < 1$. Thus $(1 - \alpha)n$ is always less than 1.

Substituting the value of n given by equation (15) into equation (16)

$$\mu = \frac{2a}{2 + \alpha l}$$

But a is the actual rate of evaporation which would prevail if the walls of the crucible were not present, i.e. if no molecules were ever to hit the surface again after emission. So $a = \mu_{max}$, where μ_{max} is the maximum rate of evaporation.

Therefore $$\frac{\mu}{\mu_{max}} = \frac{2}{2+\alpha l} = \frac{1}{1+\frac{1}{2}\alpha l} \tag{17}$$

where l is expressed in terms of multiples of the length of the radius of the cylinder.

From the molecular streaming formula due to Knudsen[4] it is easily possible to show that, for a tube of length l open at both ends

$$\frac{\mu}{\mu_{max}} = \frac{1}{1+3l/8} \tag{18}$$

This should compare with the case where no molecule is reflected from the crucible base, corresponding to $\alpha = 1$ in equations (17).

Therefore $$\frac{\mu}{\mu_{max}} = \frac{1}{1+\frac{1}{2}l}$$

Clausing[8] shows that a value of $\frac{1}{2}$ is to be expected for short tubes and 3/8 for long tubes.

It might at first sight seem unnecessary to arrive at equations (17) by the analysis given, presuming that Knudsen's well-known molecular streaming formulae could have been used. However, this analysis is essential since Knudsen's method gives no indication as to the effect of the condensation coefficient α on the final formula, unless α is put equal to 1. It could not otherwise have been foreseen that the effect of $\alpha \neq 1$ is so simple a modification in this case.

In the radiation case, where $\beta \neq 0$ and $b \neq 0$ in equation (13), the final result for the radiation from the base and walls of a short cylinder agrees precisely with the results of Buckley[3] who obtained an expression by a different method. The analysis given here is therefore capable of treating both the molecular and the heat radiation cases. The molecular case has been considered analytically on its own by Whitman,[7] who based his method on calculations made by Clausing.[8] The combination of Whitman's and Clausing's papers has some resemblance to the attack on the problem given here, but there is little doubt that the Whitman–Clausing analysis is inelegant, and in any case, not so general as the present method.

EXPERIMENTAL RESULTS USING MERCURY

To check the method of using carbon crucibles in studies of evaporation rates to investigate the validity of the formulae (17) and to arrive at a figure for α, the condensation coefficient, experiments were done on the evaporation of mercury in a vacuum from the crucibles used. Mercury was chosen because much work has been done on the determination of its vapour pressure. Temperatures were measured using a copper–eureka thermocouple inserted into the crucible base.

Table 1 shows the range of temperature and values of l/a used. The depth of the crucible l down to the mercury surface was determined from a knowledge of the mass of mercury in the crucible, the mass of mercury needed to fill the crucible, and the crucible radius, a.

Owing to the curvature of the metal surface caused by surface tension, its exact area of evaporation was unknown. If μ' represents the rate of evaporation from the crucible when the area of evaporation is taken to be equal to the cross-section area of the crucible, μ and μ' will be related by

$$\mu' = k\mu \tag{19}$$

where k is a constant giving the ratio of the area of the metal surface to the cross-sectional area of the crucible.

Since the condensation coefficient α is unknown, it is impossible to calculate μ_{max} from equation (5). So let

$$\mu_{max} = \alpha\mu'_{max} \tag{20}$$

Then $\mu'_{max} = P\sqrt{(M/2\pi RT)}$, and can be calculated from vapour pressure data.

Substituting in equations (17), using equations (19) and (20),

$$\frac{\mu'}{\mu'_{max}} = k\frac{2\alpha}{2 + \alpha l/a}$$

Rearranging gives

$$\frac{\mu'}{\mu'_{max}} = -\frac{\alpha}{2}\left(\frac{\mu'}{\mu'_{max}}\frac{l}{a}\right) + \alpha k \qquad (21)$$

Table 1. *Rates of evaporation of mercury from crucibles of depth l and radius a*

$\frac{l}{a}$	Temperature °C	$\frac{\mu'}{\mu'_{max}}$	$\frac{\mu'}{\mu'_{max}}\frac{l}{a}$
4·90	56·5	0·250	1·225
3·25	58	0·264	0·860
0·15	58	0·707	0·106
0·55	55	0·668	0·367
4·15	56	0·360	1·49
6·35	60	0·229	1·45
2·85	56	0·433	1·24
2·60	56	0·288	0·749
0·50	55	0·500	0·250
1·25	55	0·476	0·595
4·35	61	0·265	1·155
6·75	64	0·206	1·395
3·95	59	0·318	1·255
1·60	59	0·398	0·637
1·70	58	0·449	0·765

Thus a straight line should be obtained when μ'/μ'_{max} is plotted against $(\mu'/\mu'_{max})(l/a)$ using the results of Table 1. The condensation coefficient can then be calculated from the slope of the line. Once the condensation coefficient is known the rate of evaporation can be calculated with certainty from equation (5), knowing the vapour pressure P.

Fig. 5 shows the graph obtained for mercury. It will be seen that there is considerable inconsistency in the results. This is probably due to slight contamination of the mercury surface with carbon dust. (Knudsen found that for best results, using his effusion method, he had to create a new surface as frequently as possible.) Analysing the graphed results by the method of least squares, the best value for the condensation coefficient of mercury was found to be 0·7. Most workers have found a condensation coefficient of very nearly 1 for mercury.

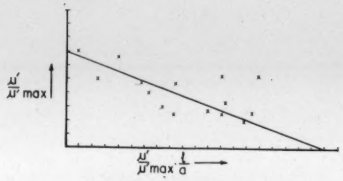


Fig. 5. Experimental results for mercury:

$$\frac{\mu'}{\mu'_{max}} \text{ versus } \left(\frac{\mu'}{\mu'_{max}}\frac{l}{a}\right)$$

An interesting point about this method is that it measures the condensation coefficient not of the saturated vapour, as is done by comparing the two sides of equation (5), but at various pressures of the unsaturated vapour. Thus, it could also show whether the condensation coefficient is dependent on the pressure of the incident vapour.

EXPERIMENTAL RESULTS FOR SILVER AND GOLD

Forty-seven experimental determinations for silver and fifty-three results for gold, both of which liquefy at elevated temperatures, proved too inconsistent to make an estimate of the condensation coefficient, and hence rate of evaporation, in the way indicated. This inconsistency is chiefly due to the difficulty of keeping the metal surface clean and of measuring the temperature with sufficient precision using an optical pyrometer. Thus an error of 5° C in the measurement of a temperature of 1500° C might well involve an error of the order of 25% in the calculation of the vapour pressure, since vapour pressures vary roughly logarithmically with temperature.

Some figures are, however, given in Table 2 (silver) and Table 3 (gold), for the approximate rates of evaporation of these metals while in the liquid state, since no experimentally obtained figures have been published hitherto. All figures are corrected to zero crucible lengths. These tables show that

Table 2. *Condensation coefficients for liquid silver*

Temperature range in °C	1000–1100	1100–1200	1200–1300	1300–1400	1400–1500
Number of determinations made	5	9	29	3	1
Maximum rate of evaporation (in g.cm^{-2} sec^{-1} × 10^{-4})	1·16	2·27	34·10	18·00	29·6
Minimum rate of evaporation (in g.cm^{-2} sec^{-1} × 10^{-4})	0·002	0·10	0·24	6·97	29·6
Mean rate of evaporation (in g.cm^{-2} sec^{-1} × 10^{-4})	0·28	0·88	5·74	11·17	29·6
Mean value of condensation coefficient	0·17	0·10	0·12	0·10	0·09

Table 3. *Condensation coefficients for liquid gold*

Temperature range in °C	1200–1300	1300–1400	1400–1500	1500–1600	1600–1700
Number of determinations made	1	9	29	13	1
Maximum rate of evaporation (in g.cm^{-2} sec^{-1} × 10^{-4})	0·08	0·35	3·41	5·39	0·90
Minimum rate of evaporation (in g.cm^{-2} sec^{-1} × 10^{-4})	0·08	0·08	0·21	0·40	0·90
Mean rate of evaporation (in g.cm^{-2} sec^{-1} × 10^{-4})	0·08	0·17	0·59	1·74	0·90
Mean value of condensation coefficient	0·68	0·35	0·43	0·40	0·08

the condensation coefficient for liquid silver in the range 1000–1500° C is about 0·1, and for liquid gold in the range 1200–1600° C is about 0·4.

ACKNOWLEDGEMENTS

The authors desire to acknowledge the assistance of Dr. M. Pirani in connexion with the use of carbon crucibles. Their thanks are also due to Dr. J. Topping in whose department this work was carried out.

REFERENCES

(1) DUSHMAN. *Scientific Foundations of Vacuum Technique* (Chapman and Hall Ltd., 1949).
(2) BUCKLEY. *Phil. Mag.*, **4**, p. 753 (1927).
(3) BUCKLEY. *Phil. Mag.*, **17**, p. 577 (1934).
(4) KNUDSEN. *Ann. Phys.* [*Leipzig*], **28**, p. 75 (1909).
(5) BARTLETT. *Phil. Mag.*, **60**, p. 111 (1920).
(6) WALSH. *Proc. Roy. Soc.*, **32**, p. 59 (1920).
(7) WHITMAN. *J. Chem. Phys.*, **20**, p. 161 (1952).
(8) CLAUSING. *Ann. Phys.* [*Leipzig*], **12**, p. 961 (1932).

Characteristics of radioluminescence in crystals

By G. T. WRIGHT, Ph.D., and G. F. J. GARLICK, Ph.D., F.Inst.P., Department of Physics, University of Birmingham

[*Paper first received* 28 *April, and in final form* 15 *June*, 1953]

Measurements of scintillation light output *versus* α-particle energy are made for single crystals of organic and inorganic substances excited by monoenergetic α-particles. A new method is used to deduce a relation between the differential efficiency dL/dx and the specific energy loss dE/dx, which in the case of calcium fluoride is verified experimentally using thin evaporated layers. The effect of phosphorescent decay is observed by varying the circuit time constant.

The intrinsic spread in scintillation pulse magnitude due to the phosphor is studied by inserting neutral filters between phosphor and photomultiplier and by variation in particle energy. Apart from a spread due to surface irregularities sources of spread are observed in diamond probably due to alternation of fluorescent and non-fluorescent crystal domains, and in anthracene, most likely due to knock-on protons near the end of the α-particle track.

The intensity of luminescence emitted by a phosphor when bombarded by nuclear particles depends on the nature and energy of the exciting particle and on the particular mechanisms of energy transfer from the particle to the crystal ions or atoms and of subsequent conversion into emission. A precise knowledge of the luminescence characteristics under such conditions is essential to the use of scintillation counters in quantitative nuclear research and is of fundamental importance in understanding the detailed mechanisms of luminescence. Crystals of organic solids such as anthracene, stilbene, etc., and of inorganic solids such as thallium activated alkali halides, scheelite and some sulphides have been previously investigated in scintillation counter systems.[1,2,3,4] The authors have extended the measurements on such crystals and have also investigated other solids, such as diamond, fluorite and single crystals of zinc sulphide. The data presented below is divided into two sections. Section 1 deals with the relation of mean scintillation intensity to the energy of the exciting particle and the decay of the intensity with time. Section 2 is concerned with fluctuations in intensity of successive scintillations which are due to intrinsic processes in the luminescent crystal as distinct from statistical fluctuations due to the photomultiplier detecting the scintillations which have been studied previously.[5]

Experimental techniques

α-particle sources used in these investigations were prepared by electrochemical deposition of polonium from acid solution on to small, polished silver disks. For such sources nearly all the α-particles are emitted with energies of 5·30 MeV ± 0·0025 MeV. The homogeneity of the beam will, however, decrease with time owing to diffusion of the polonium atoms into the silver. The extent to which this occurs is difficult to estimate, but no source was used for a period longer than six months. In addition no effects which could be attributed to a spread in energy of the emitted particles were noticed.

The particle source and the phosphor crystal were placed at opposite ends of a small chamber, about 4 cm long, in which variation of the pressure of dry air was used to control the energy of the α-particles reaching the crystal. From the geometry of the system the energy of these particles could be computed using published range-energy data.[6] The glass base of the chamber was situated directly above the semi-transparent photocathode of an E.M.I. type 5311 photomultiplier, mounted vertically. Introduction of a film of silicone oil between the phosphor crystal and chamber base and between the photocathode window and the chamber base improved the optical efficiency of the system almost threefold. A small correction, amounting to 5% at an incident particle energy of 0·3 MeV, was applied to allow for the finite sizes of the particle source and the crystal. To simplify discussion of the results measurements given below in Section 1 are for those phosphors available as single crystals of good optical quality in the form of laminae. For some of the measurements described in Section 2 arrangements were made for insertion of calibrated neutral filters between phosphor crystal and photomultiplier cathode.

Due to the different decay times of emission from the various phosphors the time constant of the multiplier collector circuit was adjusted in each case to provide effective integration of the output pulse of electrons from the multiplier due to each scintillation. The multiplier voltage was controlled within limits of one part in ten thousand by balancing against the e.m.f. of a standard cell; transient fluctuations in the dynode potentials were avoided by fitting earthing capacities. The distributions in output pulse heights from the multiplier were measured with a single channel differential pulse analyser developed in this laboratory.[7]

SECTION 1. LUMINESCENCE-PARTICLE ENERGY RELATIONS

Experimental results

Using the above methods the mean pulse height was determined for various energies of the incident α-particles

and for several specimens of each phosphor, with the exception of diamond, only one specimen being available. These measurements are presented graphically in Fig. 1. The measurements of mean pulse heights are converted to mean numbers of photoelectrons produced per scintillation; it was not thought advisable to convert completely into numbers of photons emitted by the phosphor because of uncertainty in the quoted values for photocathode conversion efficiencies.

From visual observation of the screen of a monitor oscilloscope, it was noted that for diamond and for thallium activated potassium iodide each pulse seemed to be followed by a low intensity phosphorescence. Accordingly, for these two substances, the total current from the multiplier was measured in addition to the mean height of the voltage pulses; both these sets of measurements are illustrated in curves *A* of Figs. 1(*c*) and (*d*). The relative scales of the curves were established by comparison with the corresponding measurements for an anthracene crystal. It should be noted that for the zinc sulphide crystals the luminescence decay was too long to allow pulse techniques to be used and so for this

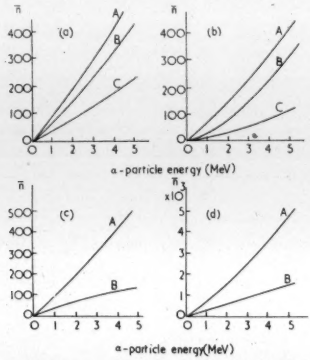


Fig. 1. Variation of mean scintillation intensity (given as mean number $\bar{n}$ of photoelectrons detected per scintillation) with α-particle energy

(*a*) *A*, silver activated zinc sulphide: *B*, anthracene; *C*, stilbene;
(*b*) *A*, barium platinocyanide; *B*, calcium tungstate; *C*, calcium fluoride.
(*c*) diamond: *A* (total current); *B* (mean number of photoelectrons $\bar{n}$).
(*d*) thallium activated potassium iodide: *A* (total current); *B* (mean number of photoelectrons $\bar{n}$).

material, the total current from the multiplier was measured as a function of incident particle energy. After completion of the measurements with α-particles the relative efficiencies of the phosphors for β-particles were determined using the conversion electrons from tin irradiated in the Harwell pile. It was assumed that all the materials gave a linear response to β-particles.

As shown by previous workers,[1,2,3,4] a more useful presentation of the information contained in the curves of Fig. 1 is given by plotting the differential luminescence output dL/dx, as a function of the differential energy loss dE/dx, along the particle path in the phosphor. The quantity dL/dE as a function of E can be obtained by direct differentiation of the illustrated curves and is shown in Fig. 2, but to proceed further requires a knowledge of dE/dx in the various crystals studied. Attempts were made to measure this quantity directly by evaporating a thin layer of the phosphor on to the surface of a calibrated, thinly aluminized, calcium fluoride crystal and measuring for the particles the energy of incidence on and emergence from the phosphor layer.

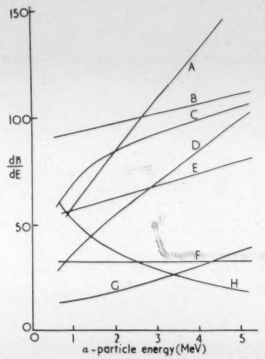


Fig. 2. Variation of $d\bar{n}/dE$ with α-particle energy for various phosphor crystals

A, thallium activated potassium iodide (total current); *B*, diamond (total current); *C*, anthracene; *D*, calcium tungstate; *E*, silver activated zinc sulphide; *F*, thallium activated potassium iodide (mean number of photoelectrons $\bar{n}$); *G*, calcium fluoride; *H*, diamond (mean number of photoelectrons $\bar{n}$).

However, sufficient accuracy of measurement could not be achieved, and eventually a semi-empirical method based on the published data of other workers was adopted.

If the results of Wilcox[8] for the stopping power of gold are replotted to show the atomic stopping power of gold relative to that of air as a function of the particle energy, a linear relation is shown for energies between 0·5 MeV and 3 MeV. This line shows no discontinuities or irregularities at particle velocities near to those of the electrons in the various atomic shells, consequently, it may be expected that for other elements a similar relation will be found. The observations of Mano[9] show that the relative atomic stopping powers of the elements increase uniformly with atomic number; there are no discontinuities or irregularities in this curve. From these two sets of measurements, and lacking any other reliable data, it was concluded that it may be admissible to express the relative atomic stopping power of any element for low α-particle energies as a straight line, the slope of which could be determined from the position of the element on the stopping power curve established by Mano. The stopping powers of compounds are then obtained by the use of the Bragg additive law. (There is some doubt, however, about the validity of this law for some substances.[10]) The curves derived by this method showing the rate of loss of energy by the α-particle as a function of its energy are shown in Fig. 3. The interpolation between 3 MeV and the points at 6 MeV, based on Mano's results, was carried out using the Bohr theoretical formula as modified for low particle velocities. Using these curves together with those of Fig. 2 the differential luminescence output, dL/dx, can be plotted as a function of the differential energy loss, dE/dx, as shown in Fig. 4. These curves also include the results of

the measurements with electrons. These studies, with α-particles and electrons, do not cover a large intermediate

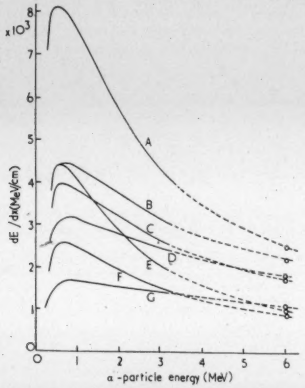


Fig. 3. Variation of specific energy loss dE/dx of α-particles with particle energy in various phosphors

A, diamond; *B*, calcium tungstate; *C*, calcium fluoride; *D*, zinc sulphide; *E*, anthracene; *F*, barium platinocyanide; *G*, thallium activated potassium iodide.

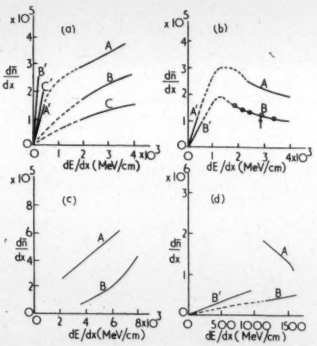


Fig. 4. Variation of specific luminescence $d\bar{n}/dx$ with specific energy loss dE/dx of α-particle for various phosphors

(*a*) *A*, zinc sulphide (αp.), *A'* (βp.); *B*, anthracene (αp.), *B'* (βp.); *C*, stilbene (αp.), *C'* (βp.);

(*b*) *A*, calcium tungstate (αp.), *A'* (βp.); *B*, calcium fluoride (αp.), *B'* (βp.);

(*c*) diamond *A* (total multiplier current) (αp.); diamond *B*, (mean number of photoelectrons $\bar{n}$) (αp.);

(*d*) thallium activated potassium iodide: *A*, total multiplier current (αp.); *B*, mean number of photoelectrons $\bar{n}$) (αp.), *B'* (βp.); (αp.) denotes α-particle excitation; (βp.) denotes β-particle excitation.

range of values for dE/dx; a suggested form which the completed curve might have is shown by the broken lines which connect the two measured ranges.

Although certain assumptions have been made in the derivation of these curves it is thought that the described procedure is better than that of previous workers who assume that the differential energy loss in the phosphor is proportional to that in air. It may be noted here that it was found possible to form thin fluorescent films of calcium fluoride by evaporation in a vacuum. Consequently, for this substance, dL/dx can be measured directly and compared with the values derived from dL/dE and dE/dx as described above. The comparison is made in Fig. 4(*b*) where the circled points denote experimentally determined values of dL/dx. These are normalized at the point indicated by the arrowhead. The good agreement obtained justifies the assumptions made in deriving the curves of Fig. 3 for this substance. Attempts were made to repeat these measurements for other materials, but although thin films could be deposited in some cases, these were non-fluorescent.

The experimental studies described above show that for diamond and potassium iodide activated by thallium the nature of the response to α-particles depends upon whether a short or a long integration time of the scintillation pulse is used. It was therefore thought desirable to examine the nature of the decay of the luminescence emission from these two materials; these measurements were subsequently extended to other phosphors.

The technique adopted consisted of determining the mean magnitude of the voltage pulses from the multiplier for various values of the time constant, (RC) of the multiplier collector circuit. Then if the decay of luminescence from the phosphor can be described by:

$$L = L_0 \exp(-\alpha t) \tag{1}$$

and if $\bar{V}$ denotes the mean pulse height, then:

$$\bar{V} = \bar{V}_0(\alpha RC)^{(1-\alpha RC)^{-1}} \tag{2}$$

This curve is shown in Fig. 5 for $\alpha = 10^7$, 10^6, 10^5, together with the experimental curves for the various phosphors. These curves are all normalized to a value of unity for V_0. That for anthracene was assumed to indicate the high frequency response limit of the electronic apparatus. The observed initial mean decay times are given in Table 1.

Table 1. *Decay time of phosphorescence for various phosphors*

Phosphor	*Decay time*
Calcium tungstate (scheelite)	$8{\cdot}3 \times 10^{-6}$ sec
Calcium fluoride	$1{\cdot}9 \times 10^{-6}$ sec
Barium platinocyanide	$1{\cdot}4 \times 10^{-6}$ sec
Potassium iodide-thallium activated	$8{\cdot}4 \times 10^{-7}$ sec
Diamond	$< 3 \times 10^{-7}$ sec

Discussion of results

As a general conclusion the results show that the relation between luminescence and particle energy is not usually linear. It depends on the nature of the phosphor crystal and also on the time constant of the light detection system comprising photomultiplier and amplifier. The relation varies with time constant in a complex way dependent on the decay characteristics of the phosphor emission. Non-linearity between luminescence and particle energy is most marked at high ionization densities. In the case of organic crystals an explanation of the fall in luminescence efficiency with increasing ionization density has been given previously by Birks.[11] The occurrence of a similar effect in inorganic crystals is most likely due to a saturation of the available emission centres particularly when the latter have relatively

long life times in the excited state. This is indicated in the results for calcium fluoride by the fact that the luminescence is approximately proportional to the depth of α-particle penetration.

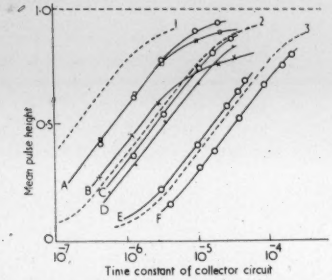


Fig. 5. Dependence of mean scintillation pulse intensity on time constant of the multiplier output circuit [see equation (2)]

Theoretical curves:
1. $\alpha = 10^7$ sec^{-1};
2. $\alpha = 10^6$ sec^{-1};
3. $\alpha = 10^5$ sec^{-1}.

Experimental curves:
A, anthracene and diamond; *B*, thallium activated potassium iodide; *C*, barium platinocyanide; *D*, calcium fluoride; *E*, calcium tungstate; *F*, neon lamp pulses.

SECTION 2. FLUCTUATIONS IN SCINTILLATION INTENSITY

Theoretical studies

Fluctuations in the output voltage pulses from the photomultiplier of a scintillation counter detecting monoenergetic particles may arise from several causes. In previous studies[5] the authors have investigated spread in pulse magnitudes due to the photomultiplier. This is mainly due to statistical fluctuations in the number of photoelectrons released from the photocathode at each scintillation. It was then shown that flawless, optically transparent phosphor crystals gave no significant contribution to the pulse spread but that imperfect specimens appeared to give a significant addition to the spread. In the following experimental and theoretical studies investigations of the sources of fluctuations in phosphor crystals of various types are made. Two main sources of fluctuations may be distinguished as follows:

i. *Optical flaws and surface conditions of the crystal.* In this case imperfections affect the penetration depth of lower energy particles and also cause a variation in the amount of light reaching the photomultiplier in successive scintillations.

ii. *Intrinsic effects.* These include variation in particle interaction with the crystal along its path for successive particles and fluctuations in the amount of energy converted into photon energy for successive particles due to the nature of the luminescence mechanism.

Methods have to be devised to distinguish between these several effects and between them and later effects occurring in the photomultiplier. Such methods have been partly indicated by the following theoretical derivations.

From previous studies a formula for the spread in pulse magnitude may be derived which is as follows:

$$\eta^2 = \eta_0^2 + \frac{5{\cdot}56}{\bar{n}}\left(\frac{s}{s-1} - pf\right) \qquad (3)$$

where η is the width of the distribution at half its maximum amplitude divided by the mean pulse magnitude $\bar{n}$ ($\eta = 2{\cdot}36 \times$ fractional variance of the distribution), η_0 is the contribution to spread due to intrinsic effects in the phosphor and effects in the light collecting system, $\bar{n}$ is the mean number of photoelectrons produced per scintillation which reach the first multiplying stage, s is the mean gain per stage, p is the probability that a photon produces a photoelectron which reaches the first dynode and f is the fraction of light emitted by the phosphor crystal which reaches the photocathode.

When there is no intrinsic effect in the phosphor and only statistical fluctuations occur then:

$$\eta^2 = \frac{5{\cdot}56}{\bar{n}}pf + \frac{5{\cdot}56}{\bar{n}}\left(\frac{s}{s-1} - pf\right) = \frac{5{\cdot}56}{\bar{n}}\left(\frac{s}{s-1}\right) \qquad (4)$$

In the case where η_0 is not zero, i.e. intrinsic spread occurs in the phosphor, the value of η_0 is obtained by inserting neutral filters between phosphor and photomultiplier to vary $\bar{n}$. For this case

$$\eta^2 = \eta_0^2 + \frac{5{\cdot}56}{\bar{n}}\left(\frac{s}{s-1}\right) \qquad (5)$$

and thus a plot of η^2 against $1/\bar{n}$ gives an intercept on the ordinate equal to η_0^2. By varying the particle energy its effect on η_0 may also be studied.

Experimental studies

(*a*) *Statistical fluctuations.* Statistical fluctuations may be determined separately by using the neon lamp source, described previously,[5] which gives recurrent pulse of light of equal intensity. By using neutral filters the validity of equation (4) may be tested. The straight lines denoted "neon lamp pulses" were obtained in this way. By comparison with these lines the pulse characteristics for phosphor scintillations may be assessed.

(*b*) *Observations for crystals with constant η_0 values.* If the particle energy is fixed at its highest value (5 MeV) and $\bar{n}$ varied by insertion of neutral filters between phosphor and photomultiplier then results are obtained as shown in Fig. 6 for various crystals. Equation (5) should be applicable to these curves, the latter giving a constant intercept η_0^2 on the ordinate. The values of η_0 thus obtained are given below:

Table 2. *Intrinsic pulse size spread η_0 for various phosphors*

Crystal	η_0
Calcium tungstate	0·26
Diamond	0·17
Thallium activated potassium iodide	0·12
Anthracene	0·00

Although a linear relation between η^2 and $1/\bar{n}$ is obtained for these crystals when $\bar{n}$ is varied by use of neutral filters the line slopes are not equal to the slope for the neon lamp pulses, with the exception of the results for anthracene which coincide with those for the neon source. These differences in slope are not easily related to any physical effect, but the following observations are perhaps relevant. It was found that using the neon lamp pulse source and obscuring the centre of the photomultiplier cathode with an opaque disk, a graph for η^2 was obtained which was 30% greater in slope than that for the case when the cathode was uniformly illuminated by the neon pulses. As the crystals were in the form of small blocks, with the exception of anthracene which was in the form of a thin lamina, geometrical shape may

influence the distribution of photons over the cathode in successive scintillations thus giving some of the observed spread in pulse magnitude.

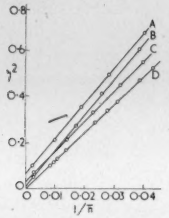


Fig. 6. Variation of pulse distribution width η with mean pulse size $\bar{n}$ for various crystals, using neutral filters to vary $\bar{n}$

A, calcium tungstate; *B*, thallium activated potassium iodide; *C*, diamond; *D*, anthracene.

(*c*) *Dependence of pulse spread* η_0 *on particle energy.* Variation in energy of the particles incident on the crystal will be likely to yield the most useful information about pulse size spread arising in the crystal. In one case the effect of surface conditions was studied for a fluorite crystal. Relations between η^2 and $1/\bar{n}$ were obtained for a freshly cleaved specimen of lamina form and then for the same crystal after the surface receiving the particle beam had been ground with fine carborundum grit (6F size). Fig. 7 shows the

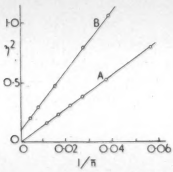


Fig. 7. Pulse spread—mean pulse size characteristics for a calcium fluoride crystal

A, freshly cleaved crystal; *B*, crystal surface ground rough.

variation of η^2 with $1/\bar{n}$ (and so with particle energy) for each case. It is easily seen that the grinding introduces an intrinsic pulse spread η_0 and the slope of the straight line graph is steeper for the ground crystal. It is likely that the grinding affects the amount of light scattered on to the photomultiplier cathode in successive scintillations. In order to obtain optimum light collection from fluorite and other crystals each specimen was mounted flush in one surface of a disc of perspex by polymerizing the latter in a mould around the crystals. The disks were attached to the photomultiplier window by a layer of silicone oil. The results for diamond and fluorite mounted in this way and for an unmounted but perfect crystal of anthracene are shown in Fig. 8. Each diagram contains the η^2 *versus* $1/\bar{n}$ relation for the neon lamp pulses. As the energy of the incident particles decreases and so $\bar{n}$ decreases there is a marked increase in η_0 for anthracene and diamond crystals. In the case of anthracene η_0 varies as $1/\bar{n}$ after a certain value of $1/\bar{n}$ is reached but for lower values it is zero. For diamond η_0^2 is proportional to $1/\bar{n}$. In the case of fluorite there is no significant pulse spread intrinsic to the crystal for the particle energy range studied.

In all above studies using α-particle excitation the density of excitation in the crystal is very high (see Section 1). To investigate the crystal behaviour for similar excitation energies, but for lower excitation densities, the particle source was replaced by a small argon discharge lamp connected in a saw tooth oscillator circuit and behaving as a source emitting pulses of ultra-violet light of constant intensity. These

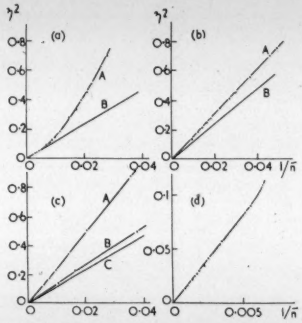


Fig. 8. Variation of pulse size spread with exciting α-particle energy (represented by $\bar{n}$)

(*a*) anthracene: *A*, particle excitation; *B*, pulses from neon lamp and "fluorescence" pulses (see text).
(*b*) calcium fluoride: *A*, particle excitation; *B*, neon lamp pulses.
(*c*) diamond: *A*, particle excitation; *B*, "fluorescence" pulses (see text); *C*, neon lamp pulses.
(*d*) anthracene: behaviour for larger values of $\bar{n}$ shown in expanded form.

pulses were filtered by a piece of Wood's glass thus giving ultra-violet light of 3500–4000 Å incident on the phosphor crystals. The fluorescence pulses excited by these pulses in the crystals were then picked up by the photomultiplier, any ultra-violet light being prevented from reaching the latter by interposing a Wratten filter between phosphor and photomultiplier. The results denoted by "fluorescence pulses" in the curves of Fig. 7 show that no intrinsic spread η_0 occurs when the density of excitation is small. Similar absence of the spread η_0 would be expected for excitation by β-particles greater than a few keV in energy, the relative excitation density being small compared with that for α-particle excitation.

(*d*) *Effect of crystal deterioration on pulse size fluctuations.* Organic scintillators, such as anthracene, stilbene, etc., show a relatively rapid deterioration under α-particle bombardment compared with inorganic phosphors. As such deterioration seriously affects their luminescence efficiencies it is useful also to investigate its effect on pulse size fluctuations. An anthracene crystal initially of good quality was subjected to prolonged α-particle bombardment. At given stages in the exposure the dependence of the intrinsic half width η_0 on mean pulse height $\bar{n}$ was measured. Curves thus obtained of η_0 *versus* $1/\bar{n}$ are given in Fig. 9. While the linear relation between η_0 and $1/\bar{n}$ is not affected there is a progressive shift in the intersect of the graph line with the abscissa showing that the intrinsic spread η_0 appears at lower pulse height as dosage of α-particles is increased. However, if we consider also the 4-fold decrease in luminescence efficiency for a dosage of 5×10^{10} α-particles/cm^2, which means that a given $\bar{n}$ value after long bombardment involves a

higher particle energy, then the effect of prolonged particle bombardment on the pulse size fluctuations appears to be small.

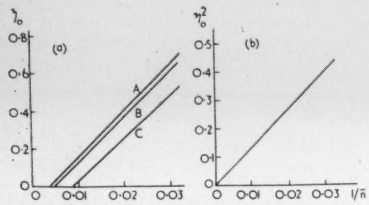


Fig. 9. Variation of intrinsic pulse spread with α-particle energy for anthracene and diamond

(*a*) anthracene; *A*, after particle dosage of zero, $1 \cdot 93 \times 10^9$, $6 \cdot 3 \times 10^9$ α/cm²; *B*, after particle dosage of $1 \cdot 6 \times 10^{10}$ α/cm²; *C*, after particle dosage of 5×10^{10} α/cm².
(*b*) diamond.

Discussion of results and conclusions

The above results show that in the case of properly cleaved calcium fluoride crystals (and for anthracene crystals excited by α-particles above a certain energy) there is no intrinsic contribution of the phosphor excitation processes to the spread in output pulse magnitude. However, for other crystals such intrinsic spread is observed and may depend on particle energy. The two cases of greatest interest in this respect are diamond and anthracene.

(*a*) *Origin of intrinsic pulse spread in diamond.* The specimen of diamond was selected because of its relatively high scintillation efficiency. Its physical characteristics are intermediate between the two well-known types of diamond distinguished in the work of Robertson, Fox and Martin.[12] It is probable that its detailed structure consists of crystal domains some of which respond to particle excitation giving luminescence emission and some of which dissipate the particle energy non-radiatively. The inference to be drawn from this structure is that an α-particle will pass through several domains before reaching the end of its path and the resulting scintillation intensity will depend on the number of luminescent domains traversed and the number of photons emitted per domain. Consider a particular case for 3 MeV α-particles incident on the diamond. The penetration depth is about 7 μ. Scintillations from such a particle produce about 100 photoelectrons which represent $\bar{X} = 5000$, the mean number of photons emitted in the crystal. From Fig. 4(*b*) and for $\bar{n} = 0 \cdot 01$ we see that η_0^2 is about 0·135, which represents for a Poisson distribution a mean number of events equal to about 40. If $\bar{S}$ is the mean number of photons emitted per domain then for such a distribution $\eta_0^2 = 5 \cdot 56\, \bar{S}/\bar{X}$ and so we obtain from the given values and for a 3-MeV α-particle $\bar{S} \simeq 120$ photons per domain. From the mean number of events equal to 40 and from the α-particle path length the size of each domain will be of the order of 10^{-5} cm which compares favourably with the probable size of mosaic blocks in the crystals.

(*b*) *Origin of pulse spread in anthracene.* The onset of an intrinsic spread η_0 for anthracene crystals as the incident particle energy is reduced at a fairly definite energy value and the linear relation between η_0 and $1/\bar{n}$, a very different behaviour from that in diamond, has to be explained. Results indicate that the spread is associated with events near the end of the particle path. In such a hydrogenous crystal there is a significant chance of an α-particle picking up electrons and thus losing its ionizing power and thus decreasing the probability of producing scintillations but increasing that of making elastic collisions resulting in knock-on protons. The latter will be able to produce ionization and, as is already known, are more efficient than α-particles of the same energy in producing scintillations in organic crystals.[13] Seitz[14] estimates that each α-particle will give rise to about thirty knock-on protons. Results in Fig. 4(*a*) show that η_0 is primarily dependent on the mean intensity of scintillations (as represented by $\bar{n}$) and not on the depth of particle penetration since a heavily bombarded crystal gives about the same value of η_0 for a given value of $\bar{n}$. The lowest useful energy of incident α-particles is 300 keV for which case η_0 is 0·6. This corresponds to a mean number of events equal to 15 or 16. For such energy each particle produces about 470 photons in the crystal and so each event produces about 30 photons. Such behaviour is most easily explained by the variation in number of knock-on protons produced by successive α-particles.

The above results and discussion indicate the necessity of avoiding phosphor crystals which show intrinsic pulse spread effects for the given particle energy range since these effects severely limit the resolution of the scintillation counter when used as an energy spectrometer for particles. Anthracene, for example, is extensively used as a scintillator though its intrinsic pulse spread η_0 is evident at α-particle energies below about 3 MeV. In the absence of intrinsic spread it is readily seen that the energy resolution is dependent only on the statistical fluctuations in the system. These can be further reduced only by the production of more efficient phosphors and more efficient cathode surfaces in photomultipliers.

As a further fundamental study it would be of interest to extend the methods used above to electron excitation of the same phosphors in which case the path lengths in phosphor crystals would be much greater than for α-particles and knock-on proton effects would be absent. In diamond, η_0 would be much reduced because of the larger number of domains traversed by each β-particle.

REFERENCES

(1) Hopkins, J. I. *Phys. Rev.*, **75**, p. 983 (1949).
(2) Franzen, W., Peelle, R. W., and Sherr, R. *Phys. Rev.*, **79**, p. 742 (1950).
(3) Frey, H. B., Grim, W. M., Preston, W. M., and Gray, T. S. *Phys. Rev.*, **82**, p. 372 (1951).
(4) Taylor, C. J., Jentschke, W. K., Remley, M. E., Eby, F. S., and Kruger, P. G. *Phys. Rev.*, **84**, p. 1034 (1951).
(5) Garlick, G. F. J., and Wright, G. T. *Proc. Phys. Soc.* [*London*], *B*, **65**, p. 415 (1952).
(6) Bethe, H. A. *Rev. Mod. Phys.*, **22**, p. 212 (1950).
(7) Wright, G. T. *J. Sci. Instrum.*, **29**, p. 157 (1952).
(8) Wilcox, H. A. *Phys. Rev.*, **74**, p. 1743 (1948).
(9) Mano, G. *Ann. Phys.* [*Paris*] (11th Ser.), **1**, p. 407 (1934).
(10) Appleyard, R. K. *Proc. Cambridge Phil. Soc.*, **47**, (Pt. 2), p. 443 (1951).
(11) Birks, J. B. *Phys. Rev.*, **84**, p. 364 (1951).
(12) Robertson, R., Fox, J. V., and Martin, A. E. *Phil. Trans. A*, **232**, p. 463 (1934).
(13) Birks, J. B. *Proc. Phys. Soc.* [*London*] *A*, **64**, p. 874 (1951).
(14) Seitz, F. *Disc. Faraday Soc.* (No. 5), p. 271 (1949).

Line interaction in X-ray and electron diffraction photographs

By G. D. Archard, A.Inst.P.,* The University, Reading

[*Paper first received* 17 *April and in final form* 6 *July*, 1953]

The mutual effect of two neighbouring lines on X-ray and electron diffraction patterns is treated theoretically. The theory is in qualitative agreement with the experimental results of electron diffraction photographs (taken on *plates*) but cannot be used to correct these quantitatively on account of the practical difficulty of making the necessary measurements. In certain X-ray photographs (taken on *films*) there arises an effect, not treated by the theory, which for low values of Bragg angle acts in an opposite sense. The general conclusion is that for precision measurements close groups of lines should be treated with circumspection or altogether avoided.

It has been shown by Rymer and Butler[1] that the measurement of electron diffraction ring radii is subject to certain errors, in particular that due to the presence of diffuse background. They gave an expression for its correction in the form:

$$\text{Background error} = (k/I_l)\Delta I_b W \qquad (1)$$

where I_l was the peak intensity of the line, ΔI_b was the change in background intensity over the width W of the line, W was the distance between points of intensity $I_l/2$, and k was a constant depending on the geometrical shape of the line. The sense of the correction could be determined from elementary considerations.

Now frequently in electron diffraction photographs two lines fall so closely together that the flank of one forms an extra background to the other. In X-ray diffraction, the same phenomenon is more widely experienced when radiation involving doublets (e.g. $CuK\alpha_{1,2}$) is used; when the members of a doublet are only just resolved they may be expected to disturb each other.

It is the purpose of the present paper to investigate theoretically the mutual effect of neighbouring lines and to examine how far the theory may be considered to correspond to experimental facts.

Mutual Background

Fig. 1 represents a hypothetical case of line interaction. The trace revealed by a linear microphotometer is shown as a thick line; dotted lines represent the components. If the lines are known to be symmetrical (e.g. Gaussian, which is a reasonable approximation to the conditions found in practice), analysis could be effected by reflecting the unaffected sides in the ordinates through the peaks; this is a well-known principle and its particular application will now be considered.

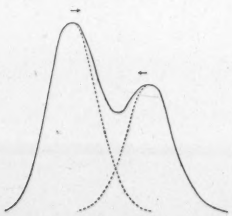


Fig. 1. Synthesis of two lines of Gaussian form

As a first step it will be assumed that the microphotometer slit is infinitesimal; corrections for the effect of finite slits can be made but will not be required here. It will also be assumed that the blackening of the photograph is proportional to the intensity I. To increase generality, a triplet with diffuse background will be considered. This may be expressed:

$$I = mx + A_1f_1(x^2) + A_2f_2(x - x_2)^2 + A_3f_3(x - x_3)^2 \qquad (2)$$

where x represents the radial distance from the centre of the first line of the triplet, and the background, assumed to be linear, is represented by mx. Representation as $f(x^2)$ instead of $f(x)$ accounts for the line symmetry.

Peaks exist where:

$$\tfrac{1}{2}m + A_1xf_1'(x^2) + A_2(x - x_2)f_2'(x - x_2)^2 + A_3(x - x_3)f_3'(x - x_3)^2 = 0 \qquad (3)$$

where the dash indicates differentiation with respect to x.

One peak will be *near* $x = 0$, and this will be considered. Expand in the form:

$$f_1(x^2) = 1 + a_1x^2 + b_1x^4 \ldots \qquad (4)$$

$$f_2(x - x_2)^2 = 1 + a_2(x - x_2)^2 + b_2(x - x_2)^4 \ldots \qquad (5)$$

$$f_3(x - x_3)^2 = 1 + a_3(x - x_3)^2 + b_3(x - x_3)^4 \ldots \qquad (6)$$

Substitute equations (4), (5), and (6) in equation (3), neglect powers of x above the first, and after a little analysis deduce:

$$x = -\frac{m + (slope_2)_0 + (slope_3)_0}{2A_1a_1 + d/dx\,.\,(slope_2)_0 + d/dx\,.\,(slope_3)_0} \qquad (7)$$

(where $(slope_2)_0$ means slope of line 2 at $x = 0$).

This gives the general background correction for the line with subscript (1); in any computation it would be arranged that the line under investigation should have this subscript. On the assumption of the line being of Gaussian form, the quantity a_1 would become $-1/2\sigma_1^2$. The expression (7) would then resemble:

$$x = m/(A_1/\sigma_1^2) \qquad (8)$$

with correction terms in numerator and denominator; equation (8) is merely equation (1) with a change of notation.

Consider equation (7) in relation to the middle line of a close triplet (e.g. the 100, 002, 101 of a hexagonal lattice of the zinc oxide form, Fig. 3). If the combined trace be analysed geometrically as described above, the slopes of the 101 and 100 under the peak of the 002 may be measured, added to the slope of the diffuse background, and the correction deduced. However, the slopes of the flanking lines will be of opposite sign and similar order, so that in order to determine their sum each must be known with considerably more accuracy than is necessary in the final correction term.

In practice this is found a major disadvantage, for the geometrical analysis of the lines is itself somewhat approximate, and the determination of the slopes of the individual lines introduces further error. In consequence it is only

* Now at Associated Electrical Industries Ltd., Aldermaston, Berkshire.

rarely that the expression (7) can be of practical value. It does, however, demonstrate the danger of using conflicting lines in the determination of crystal structure, as will appear in the following two sections.

AN EXAMPLE OF LINE INTERACTION IN ELECTRON DIFFRACTION

Rymer and Butler[1] showed that due to other causes the values of $\lambda L/a$ determined from various rings (λ = electron wavelength, L = camera length, a = lattice constant) would be expected to be a linear function of $1/R^2$, where R = ring radius. Fig. 2 shows the relationship actually obtained for

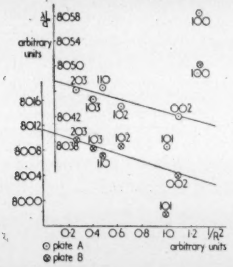


Fig. 2. Values of $\lambda L/a$ plotted against $1/R^2$. Showing the difference in lattice spacings a deduced from the close triplet 100–002–101 and other diffraction rings on two electron diffraction photographs of zinc oxide (axial ratio assumed 1·60220)

two typical photographs of zinc oxide in which the effect of line interaction had not been considered and the background corrections applied were merely those pertaining to the diffuse background. The members of the triplet 100, 002, 101 appear to possess very different lattice constants, but a little consideration will show that the discrepancies are of the sign which would be expected on the assumption that the flanks of the 002 draw the peaks of the 100, 101 inwards, while the flanks of the 100 and 101, together with the diffuse background, very nearly produce a level total background behind the peak of the 002 (see Fig. 3). Smaller effects of the same nature are found for other close lines.

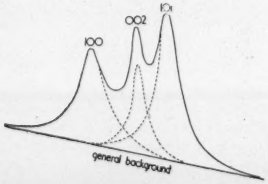


Fig. 3. Microphotometer trace of the 100–002–101 group of rings in an electron diffraction photograph of zinc oxide. This shows the uncertain effect of the flanks of lines on the positions of neighbouring peaks

If the axial ratio of zinc oxide be calculated from the first six rings (100, 002, 101, 102, 110, 103) of photographs such as those indicated, its value appears to be of the order of 1·605, which is almost comparable with the value 1·607 given by Finch and Wilman.[2] On the other hand, when only the 102, 110, and 103 rings were used the mean axial ratio found from eight plates was 1·60220 ± 0·00003, which compares favourably with two sets of X-rays values obtained by Rymer and Archard,[3] respectively 1·602195 ± 0·000005 and 1·60220 ± 0·00004.

Application of the expression (7) to the close triplet brings the axial ratio calculated from the six rings down in the right direction, but its effect cannot be determined sufficiently accurately to yield agreement as close as that quoted for three rings. The general conclusion in the case of electron diffraction lines is that the expression (7) represents the circumstances qualitatively, but that the quantities involved cannot be determined sufficiently accurately for the expression to be of use. Hence close groups of lines should be omitted from precision calculations.

AN EXAMPLE OF LINE INTERACTION IN X-RAY DIFFRACTION

Consider the case of a just-resolved doublet. This may be represented by omitting m and quantities with the subscript $_{(3)}$ from equation (7). Then the theory indicates that the members of the doublet will attract one another. As the numerator contains the background slope and the denominator contains the peak intensity, it is evident that the stronger member of a doublet will affect the weaker more than *vice versa*. If, therefore, "background" were the only mutual effect, it would be appropriate to assign always greater weight to the stronger line, quite apart from the question (as yet controversial) as to whether it is easier to

Table 1. *Values of $\lambda/2 \sin\theta$ for doublets of a sodium chloride and a zinc oxide film*

(The α_1 is the first quoted of each pair)

	Sodium chloride			*Zinc oxide*	
hkl	$\lambda/2 \sin\theta$	*Differences*	*hkl*	$\lambda/2 \sin\theta$	*Differences*
	1·2576686			0·9354807	
210		7105	300		3417
	1·2569581			0·9351390	
	1·1483481			0·9044639	
211		659	213		1222
	1·1482822			0·9043417	
	0·9950334			0·8802651	
220		5544	302		1366
	0·9944790			0·8801285	
	0·9380862			0·8654342	
221		3428	006		− 404
	0·9377434			0·8654746	
	0·8899604			0·8350148	
310		1740	205		− 493
	0·8897864			0·8350641	
	0·8485625			0·8272061	
311		311	106		−1164
	0·8485314			0·8273225	
	0·8125395			0·8216737	
222		−1063	214		−1615
	0·8126458			0·8218352	
	0·7807029			0·8105503	
302		70	220		− 564
	0·7806959			0·8106067	
				0·7789149	
			310		− 53
				0·7789202	

sight a microscope on a strong line than on a weak line. That the simple theory is inadequate will now be demonstrated by reference to experimental results.

In a 19 cm Debye–Scherrer powder camera various specimens were irradiated and the resulting films were each measured carefully several times. The values of $\lambda/2 \sin\theta$ deduced from the various Bragg angles were computed; results of two typical films appear in Table 1.

Apart from one irregularity in the sodium chloride results, the following may be said to represent both films: for low Bragg angles the difference between the values of $\lambda/2 \sin\theta$ pertaining to the α_1 and α_2 radiations is very large; it diminishes to zero in the neighbourhood of $\theta = 60°$ and then becomes fairly large in the opposite sense. Finally, it diminishes as if to return to zero at $\theta = 90°$.

The remarkable fact is that the low angle difference corresponds to a *repulsion* instead of the *attraction* expected on the background theory. It has been mentioned[4] that during the development of two neighbouring lines chemical action at the peaks weakens the developer between the peaks and causes an apparent repulsion. This suggests that at low angles the developer effect predominates, but that the mutual background effect overtakes it as the lines separate on advancing to higher Bragg angles. Finally, as the lines separate completely the total effect tends to zero. That the observations recorded in Table 1 are no mere chance is

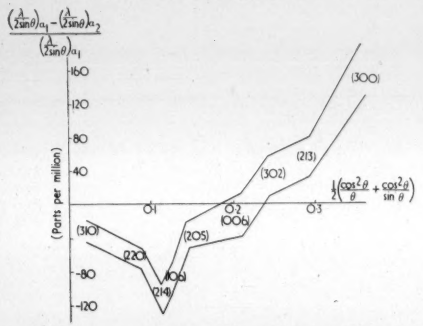


Fig. 4. Difference between values of $\lambda/2 \sin\theta$ calculated from the α_1 and α_2 members of doublets averaged over fourteen X-ray photographs of zinc oxide. The spacing of the lines indicates twice the standard deviation

attested by Figs. 4 and 5. In Fig. 4 appear the values of $(\lambda/2 \sin\theta)_{\alpha 1} - (\lambda/2 \sin\theta)_{\alpha 2}$ of a group of fourteen zinc oxide films. The distance between the lines represents twice the standard deviation taken over the fourteen films. For convenience the scale of abscissae is taken as the extrapolation function of Nelson and Riley.[5] Fig. 5 represents some of the mean extrapolation residuals taken over a similar group. (Values of lattice spacing a obtained from various lines are plotted against $\frac{1}{2}(\cos^2\theta/\theta + \cos^2\theta/\sin\theta)$. Theoretically the result should be a straight line. The *residuals* of the least squares plot represent coherent or incoherent deviations from theory.) The sharp double bend comes in the middle of a group of four doublets, the inner ones of which (106, 214) are particularly close. Here each individual line (i.e. half a doublet) may be considered to be affected by at least three others. It is reasonable to expect, moreover, that the centroids of the 106 doublet and the 214 doublet will either move together or move apart. The sense of the double bend in the diagram indicates that the doublets *repel* each other. This is surprising because Fig. 4 indicates that in the region of these doublets individual lines *attract*. Clearly there is a

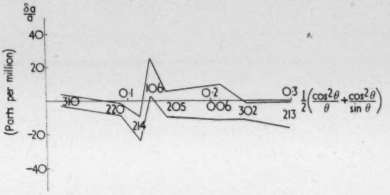


Fig. 5. Extrapolation residuals averaged over thirteen X-ray diffraction photographs of zinc oxide. The spacing of the lines indicates twice the standard deviation

complicated system of mutually induced errors in the case of such groups and, as in the case of electron diffraction plates, these might well be omitted for precision measurements.

Had only one mutual effect obtained it would certainly have been appropriate to give the stronger line greater weight in computations, but in view of the evident complexity of interaction this question may well be submitted to the test of consistency, as follows.

Table 2. *Comparison of the results of calculating the lattice constants of certain zinc oxide specimens using equal and unequal weights for the members of the CuK$\alpha_{1,2}$ doublets*

Film	Lattice spacing Weighting		Axial ratio Weighting	
	Equal	Unequal	Equal	Unequal
	kilo X units			
32	3·243283	3·243229	1·602401	1·602355
	25	55	85	83
33	3308	3284	2316	2272
44	3003	2995	1987	1941
	3	41	23	40
45	3006	2954	1964	1901
48	3031	2996	2071	2051
	3	30	17	33
49	3034	2966	2088	2084
53	3029	2991	2166	2143
	22	26	10	32
54	3007	2965	2156	2111
55	3074	3043	2136	2141
	45	12	14	39
56	3029	3031	2150	2102
57	3097	3073	1700	1653
	24	47	20	25
58	3073	3026	1720	1678
61	3051	3020	2163	2141
	5	6	40	51
64	3056	3026	2203	2192

Table 2 gives the results of certain zinc oxide photographs calculated firstly by weighting α_1 and α_2 equally, and secondly

by giving α_1 double weight. There is slight evidence that the former gives the more consistent results. It should be mentioned that pairs of films linked by braces represent two films of the same specimen. It is mentioned elsewhere[3] that the axial ratio of zinc oxide bears a linear relation to the amount of cadmium contamination. If the values of axial ratio computed with equal and unequal weighting are plotted against cadmium content and the line extrapolated to zero cadmium, the respective results are:

Equal weighting: $1{\cdot}602195 \pm 0{\cdot}000005$
Unequal weighting: $1{\cdot}602172 \pm 0{\cdot}000010$

The difference is 1 in 70000, which is only just significant.

It should perhaps be emphasized that all the results quoted above relate to regular zinc oxide. It is shown elsewhere[6] that certain samples have crystallographic abnormalities, but these are not to be confused with the photographic effects here described. Indeed, the films pertaining to the abnormal specimens were extremely clear and showed *less* line interaction than those quoted.

CONCLUSIONS

Neighbouring diffraction lines on photographic plates or films affect each other so seriously that it is best in general to refrain from using close groups in crystallographic determinations. In the case of X-ray diffraction only well-resolved doublets should be used and the members of these may then be assigned equal weights. Regions where several doublets fall close together should be treated with circumspection. The considerations here set forth may well have wider applications, for example in spectroscopy.

ACKNOWLEDGEMENTS

I wish to thank Professor R. W. Ditchburn for laboratory facilities, Dr. T. B. Rymer for helpful discussions, and the Department of Scientific and Industrial Research for a grant operative at the time when the investigations were carried out.

REFERENCES

(1) Rymer, T. B., and Butler, C. C. *Phil. Mag.*, **36**, p. 515 (1945).
(2) Finch, G. I., and Wilman, H. *J. Chem. Soc.*, **1**, p. 751 (1934).
(3) Rymer, T. B., and Archard, G. D. *Research*, **5**, p. 292 (1952).
(4) Butler, C. C. Ph.D. Thesis, University of Reading, p. 44 (1946).
(5) Nelson, J. B., and Riley, D. P. *Proc. Phys. Soc.* [*London*], **57**, p. 160 (1945).
(6) Archard, G. D. *Acta Cryst.*, **6**, p. 657 (1953).

The depression of liquid surfaces by gas jets

By R. D. Collins, B.Sc., and H. Lubanska, British Iron and Steel Research Association, London, W.1

[*Paper first received* 26 *August*, 1952, *and in final form* 19 *August*, 1953]

A gas jet directed on to a liquid surface causes considerable agitation of the liquid, a fact which is largely responsible for the rapidity of the steel-making reactions in the side-blown converter[1] as compared to the open-hearth furnace. As a start for laboratory investigation of this effect, the disturbance caused by an air jet playing on a water surface was photographed. The maximum depth of penetration was measured and plotted against the calculated reaction thrust, $\rho Q^2/A$, for a range of jet direction, orifice diameter, distance of orifice from water surface and air velocity. The consistency of the results within the practical limits of experimental error suggested an empirical relationship between the variables for angles between jet and surface above about 25°, and the following equation was derived expressing penetration as a function of thrust and angle of jet:

$$P = 53\tau \sin\theta/[x^2\omega + 19(\omega\tau^2)^{\frac{1}{3}}]$$

LIST OF SYMBOLS

P = maximum depth of depression, cm;
x = slant distance of jet orifice from undisturbed water surface, cm;
θ = angle between jet axis and projection on water surface;
d = diameter of jet orifice, cm;
Q = flow-rate of gas, c.c./sec;
A = area of jet orifice $= \pi d^2/4$ cm^2;
τ = reaction thrust $= \rho Q^2/A = (\pi/4)\rho v_0^2 d^2$ dynes;
ρ = density of gas, g/c.c.;
s = density of liquid, g/c.c.;
w = specific weight of liquid (i.e. density × gravitational constant);
v_0 = mean velocity of gas at jet mouth, cm/sec;
μ_1 = viscosity of gas, poises;
μ_2 = viscosity of liquid, poises;
σ = surface tension of liquid, dynes/cm.

INTRODUCTION

In previous work on the flow of fluids in a side-blown converter,[1] it became evident that there was a wide field for study of the effects which arise when a jet of air strikes a liquid surface. In the side-blown converter, interest is focused upon the transfer of oxygen to the bath surface by means of the jet and the subsequent diffusion of the oxygen to the lower layers of metal by the turbulent motion of the bath.

The opportunity of carrying out direct experiments on the impact of air on to the surface of a bath of molten steel has not yet presented itself, but in the meantime, work has been done on the impact of a jet on to the surface of a liquid at room temperature. It is planned that experiments shall be extended to molten metal when this becomes practicable.

AIM OF PRESENT EXPERIMENTS

As a first step to the general study of the interaction of jets and liquid, it was decided to investigate the disturbance of the surface which is brought about by the jet impact. A series of experiments was undertaken in which the depth of depression caused by an air jet directed on to a water surface was measured.

The area of the free water surface was large in comparison

with the disturbance, and tests confirmed that interactions between liquid and container could be ignored. (The ratio of the water surface to the area of disturbance would be of the order 100 : 1 in the present experiments.) The maximum depth of depression was measured for a range of impact angle, nozzle diameter, distance of nozzle from liquid and air velocity.

APPARATUS AND EXPERIMENTAL PROCEDURE

For the experimental tests a glass-sided tank of water ($2 \times 2 \times 2$ ft) was used (see Fig. 1). An adjustable jet tube was set up to point towards the centre of the liquid surface.

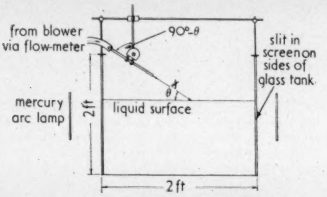


Fig. 1. Experimental arrangement

The angle of the jet tube, the distance of the orifice from the water and the diameter of the tube could be varied and measured. The tube was connected by rubber tubing to a compressor, the air-flow being measured by a Rotameter which was calibrated against an orifice.

Series of photographs were taken with a Kine Exakta camera of the depression caused by the air jet playing on the liquid surface. Slit illumination from the two sides of the screened tank by means of mercury arc lamps gave a clear picture of the water displacement. The maximum depth was measured from the 35 mm film negatives with the aid of a projector and graticule scale. This could possibly have been done more accurately from enlargements, but the time taken for each measurement would have been very much greater, and in view of the inevitable degree of scatter in this type of experiment, it seemed better to make measurements of a greater number of experiments and obtain an average picture of the phenomena, than to concentrate on extreme accuracy of individual results.

THEORETICAL TREATMENT

It does not appear possible to treat the problem analytically even when a number of simplifying assumptions have been made. However, by applying the methods of dimensional analysis the number of variables to be treated can be reduced. Further, if a number of assumptions are made which can be shown to be reasonable approximations under the conditions of the experiment, the problem is reduced to one of only three variables. By this means the presentation of results is greatly facilitated.

The flow of gas and the flow of liquid which together make up the system are not independent but are connected by the boundary conditions that pressure and shear stress are continuous across the boundary. As an introduction to the general case the simplifying assumption may be made that the pressure at the surface is equal to the hydrostatic pressure (i.e. that changes in dynamic head in the liquid are negligible). This implies that the velocity of the liquid at the surface is small compared with the velocity of the gas.

As far as the gas flow is concerned the liquid then merely provides the boundary condition that the pressure at the surface is wy where y is the virtual distance of a point on the surface below the undisturbed level, and w is the specific weight of the liquid. If P is the depth of the depression made by the jet the quantities on which it can depend are (d, v_0, x, θ, μ_1, ρ and w). Of these, the first four specify the position, size and strength of the jet, the next two the flow properties of the gas, and the last the characteristic of the boundary surface. For convenience, the quantity

$$\tau = \frac{\pi}{4}\rho v_0^2 d^2$$

the reaction thrust of the jet, may be taken as an independent variable instead of v_0.

Thus we have a relation connecting P to these seven independent variables.

$$f(P, d, \tau, x, \theta, \mu_1, \rho, w) = 0$$

By the Π theorem[2] this relation between seven variables can be expressed in terms of four dimensionless variables.

$$\phi\left[\frac{P}{(\tau/w)^{\frac{1}{3}}}, \frac{x}{(\tau/w)^{\frac{1}{3}}}, d/x, \theta, (\tau\rho)^{\frac{1}{2}}/\mu\right] = 0$$

where ϕ is an unknown function of the five variables. Or

$$\frac{P}{(\tau/w)^{\frac{1}{3}}} = \psi\left[\frac{x}{(\tau/w)^{\frac{1}{3}}}, d/x, \theta, (\tau\rho)^{\frac{1}{2}}/\mu\right]$$

Or

$$P' = \psi[x', d', \theta, R_1]$$

where $P' = P(w/\tau)^{\frac{1}{3}}$, $x' = x(w/\tau)^{\frac{1}{3}}$, $d' = d/x$, $R_1 = (\tau\rho)^{\frac{1}{2}}/\mu$, and ψ is some function of x', d', θ and R_1.

Of these dimensionless quantities, R_1 is seen to be $(\pi/4)^{\frac{1}{2}}$ times the Reynolds number of the gas at the jet mouth and $1/d'$ the distance of the surface from the mouth of the pipe in diameters. But it is known that variation of Reynolds number has little effect on the behaviour of a fully turbulent jet as far as mean conditions are concerned. Again, at points more than 8 diameters from the mouth of the jet the flow does not depend on the diameter of the mouth, but only on the thrust τ, the distance x from the source, and the density ρ.

It is clear that under these conditions P' will be determined mainly by x' and θ and only to a much lesser extent by d' and R_1.

In this development we have made the approximation that the pressure at the surface is the hydrostatic pressure which supposes that there is no dynamic pressure gradient in the liquid. If this assumption is removed it is necessary to introduce, as additional parameters, the characteristics of the fluid, μ_2, s, σ.

It is necessary to regard s and w as independent parameters since the terms in the equations of motion which correspond to buoyancy pressures are proportional to specific weight, but those corresponding to inertial forces are proportional to density.

In the same way as before we obtain

$$P' = \psi_2[x', d', \theta, R_1, R_2, s', W]$$

when

$$R_2 = \frac{1}{\mu_2}(\tau s)^{\frac{1}{2}}$$

$$s' = s/p$$

$$W = \left(\frac{w\tau^2}{\sigma}\right)^{\frac{1}{3}}$$

But it is found that within a wide range of conditions the effect of these three new parameters is small. R_2 and W are respectively the Reynolds number of the liquid motion and the Weber number of the interface, and as viscosity and surface tension are quite low it is not surprising that their effect is small.

Insensitivity to a change of s' (the ratio of gas to liquid density) can also be simply explained, for the depth of the depression will be reduced to a first approximation by the dynamic head of the liquid welling up beneath it and this in turn will be proportional to the shear stress at the surface which maintains the liquid motion. Providing the velocity of the gas is much greater than that of the liquid (s' large) this stress will be independent of the velocity in the liquid and depend only on the dynamic head of the gases. Furthermore, the dynamic head in the liquid is found to be small compared to the hydrostatic head of the depression.

Of the seven variables, those which determine the system, two, namely x' and θ, are of the greatest importance, and it remains to find experimentally a relationship $P = f(x, \theta)$ and confirm that it holds over as wide a range of all the variables as is possible within the practical limits of the simple form of experiment adopted.

RESULTS

The experiments were designed with two objects in view, to determine an empirical relationship $P' = f_3(x', \theta)$ or $P' = f_3[x(w/\tau)^{\frac{1}{3}}, \theta]$ and to confirm its independence of the other variables over as wide a range as could conveniently be arranged.

In one set of experiments a fixed axial distance, x, tube diameter, d, and angle of jet, θ, were chosen, and the rate of flow varied. With an axial distance $x = 7{\cdot}6$ cm tests were then made with values of tube diameter $d = 1{\cdot}12$, $0{\cdot}794$, $0{\cdot}715$, $0{\cdot}614$ cm, and $\theta = 90°$, $60°$, $45°$, $30°$ and $15°$. As expected, no significant variation occurred with change of the value of d (Fig. 2), and the mean values of P against τ are plotted in Fig. 3.

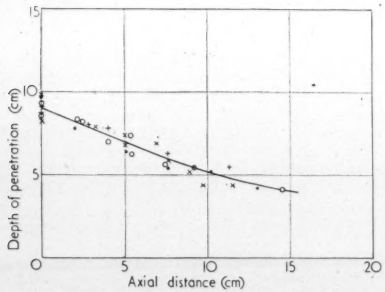


Fig. 2. Effect of axial distance with 90° impingement

Angle of jet = 45°.
Thrust $\rho Q^2/A = 43\,000$ dynes.
○ tube 1·12 cm diameter.
+ tube 0·794 cm diameter.
● tube 0·715 cm diameter.
× tube 0·614 cm diameter.

A second series of experiments was then undertaken in which the thrust was kept constant and the distance x varied. Again the mean of the results for the four tubes was taken and Fig. 4 shows a plot of P against x for the five values of θ.

In each case it is seen that the curves for the various jet angles are very similar, which suggests a relationship of the type

$$P = f_1(\theta) \cdot f_2[x(w/\tau)^{\frac{1}{3}}]$$

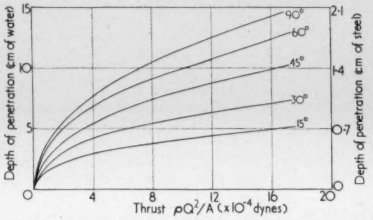


Fig. 3. Average curves for four tubes at 7·6 cm axial distance

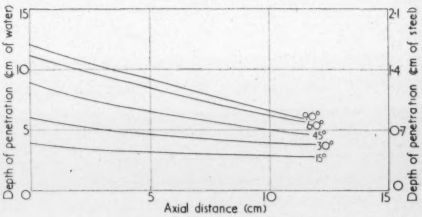


Fig. 4. Average curves for four tubes at thrust approximately 42 000 dynes

If an abscissa is chosen on Fig. 3 or 4 and the corresponding ordinates read off for the five values of θ, five points are obtained for a plot of P against θ. Fig. 5 shows two such sets of points compared with a sine curve, and the agreement is seen to be good as far as the 30° and higher angles are concerned. A number of results were obtained below the 30° angle of jet, and although the accuracy of the experiment is lower under these conditions, where the depth

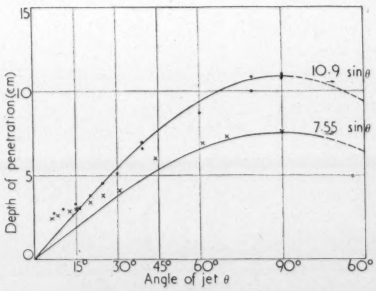


Fig. 5. Divergence of experimental points from sine curves at angles of jet below approximately 30°

Thrust = 70 000 dynes.
Tube = 0·715 cm diameter.
● axial distance = 7·6 cm.
× axial distance = 15·8 cm.

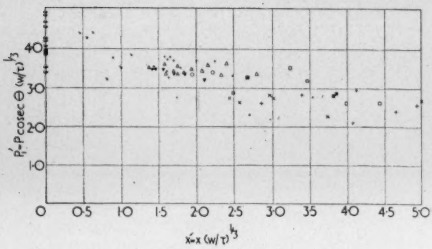


Fig. 6. Plot of points with 15° angle of impingement
Angle of jet = 15°.

of depression is small and the degree of splashing high, it is quite clear that the sine relation does not hold at small angles. Further, the shape of the curve is no longer independent of x'.

In the previous section it was shown that it should be possible to express the results in the form $P' = f(x', \theta)$. The above results were then plotted in the form of

$$P'_1 = \frac{P'}{\sin\theta} = P(w/\tau)^{\frac{1}{3}} \operatorname{cosec}\theta \text{ against } x' = x(w/\tau)^{\frac{1}{3}}$$

Figs. 7(*a*) and (*b*) show, respectively, a plot for a jet angle of 90° and a combined plot for all angles except 15°. The remarkable consistency of this latter plot in which the scatter is well within the limits of experimental error over a wide range gives strong support to the choice of variables.

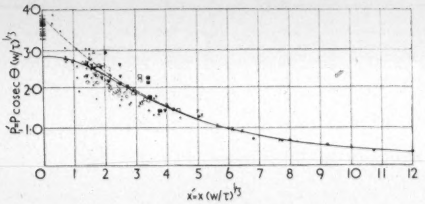


Fig. 7(*a*). Plot of points with 90° impingement

Angle of jet = 90°.
$P'_1 = 53/(x'^2 + 19)$.
$P'_1 = 53/(x'^2 + 19) + \exp(-x')$.

$x/d = 0$	□	$x/d = 14$	□
< 5	◕	$= 17$	■
$= 5$–8	●	$= 21$	+
$= 9{\cdot}6$	△	$= 25$	×
$= 10{\cdot}6$	○	$= 48$	◉
$= 12{\cdot}5$	▼	$= 12{\cdot}5$	◇ CO_2

As x' becomes large the surface of the liquid becomes more nearly smooth, and in the limit the system approaches a jet playing on a flat surface. It is to be expected, therefore, that in the limit P' should be inversely proportional to x'^2 and in the case of normal incidence the hydrostatic pressure of the depression should be of the same order as dynamic head at the same distance along the axis of a free jet. Using the formula published by Hinze,[3] the latter condition implies that for large x' $p'x'^2/26$ is of the order of unity. As the experimental limit is 53/26, extrapolation to larger values of x' should be reliable.

A curve of the form $p' = a/(x'^2 + b)$ was fitted to the data of Fig. 7(*b*) and with $a = 53$, $b = 19$, a satisfactory fit is obtained which, however, departs considerably from the experimental points for small values of x'. The dotted curve represents the equation

$$p'_1 = 53/(x'^2 + 19) + \exp(-x')$$

which fits better, but the behaviour of the system for small x' appears complicated and the experimental scatter was considered too great to warrant additional empirical constants.

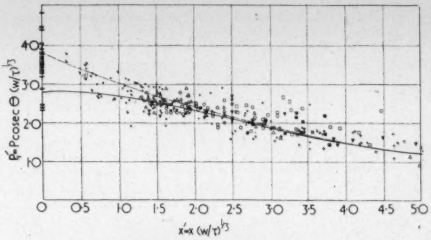


Fig. 7(*b*). Combined plot of points with 30°, 45°, 60° and 90° impingement

$P'_1 = 53/(x'^2 + 19)$.
$P'_1 = 53/(x'^2 + 19) + \exp(-x')$.

$x/d = 0$	□	$x/d = 12{\cdot}5$	▼
< 5	◕	$= 14$	□
$= 5$–8	●	$= 17$	■
$= 9{\cdot}6$	△	$= 21$	+
$= 10{\cdot}6$	○	$= 25$	×

Fig. 8 shows the results of two independent experiments in which air and carbon dioxide jets were used under similar conditions. Maximum depth of depression is plotted against

$$h(p_1/p)^{1/\gamma}$$

where h is the orifice differential, in water, p_1 is the absolute pressure, p is the atmospheric pressure, γ is the ratio of specific heat coefficients, and $h(p_1/p)^{1/\gamma}$ is therefore proportional to thrust, τ, under similar experimental conditions:

$$h(p_1/p)^{1/\gamma} = 2{\cdot}42 \times 10^{-4}\tau$$

The points obtained indicate that neglect of the density ratio parameter, s', is justifiable within the practical limits of the experiment adopted.

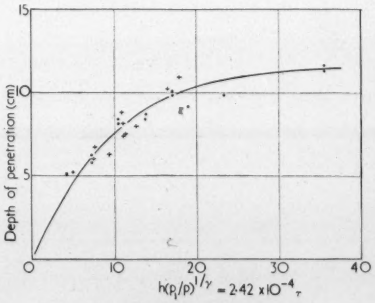
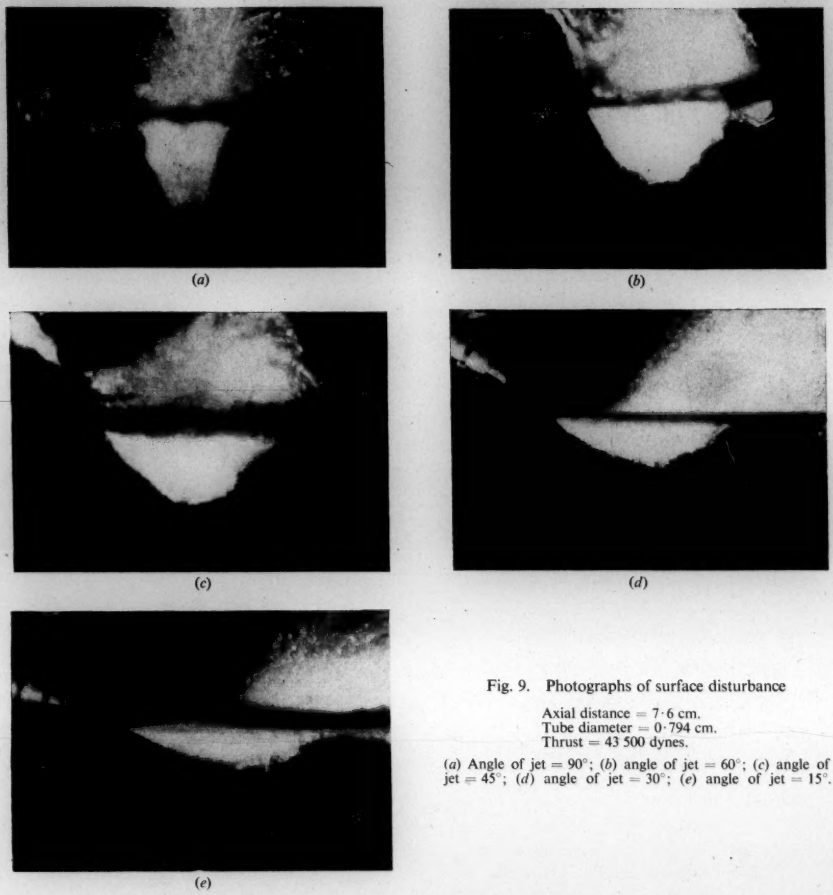


Fig. 8. Change of density ratio
Angle of jet = 90°.
Axial distance = 7·6 cm.
● air, $s/\rho = 816$. + CO_2, $s/\rho = 530$.



(*a*) (*b*) (*c*) (*d*) (*e*)

Fig. 9. Photographs of surface disturbance

Axial distance = 7·6 cm.
Tube diameter = 0·794 cm.
Thrust = 43 500 dynes.

(*a*) Angle of jet = 90°; (*b*) angle of jet = 60°; (*c*) angle of jet = 45°; (*d*) angle of jet = 30°; (*e*) angle of jet = 15°.

CONCLUSION

In the experiments described the maximum depth of depression has been taken as the principal criterion, since this was the most easily measured characteristic of the liquid disturbance. However, as the photographs (Fig. 9) show, maximum depth of depression does not necessarily imply maximum mixing and gives only limited information of the extent and character of the whole disturbance. Further measurements and calculations to determine the actual volume displacement will be carried out.

In the meantime, the present results show good agreement with the simplified theory outlined, and form a substantial basis for further research. In view of the inevitable scatter found in this type of experiment, a very large number of tests were made in order to arrive at the mean curves, from which the empirical relationship

$$P' = f\left[x(w/\tau)^{\frac{1}{3}}, \theta\right]$$

was determined. The establishment of this relationship with the derivation of an empirical equation, and the confirmation of its independence (within a large range of other variables) form the main achievement of the present project.

REFERENCES

(1) Newby, M. P. *J. Iron Steel Inst.*, **162**, p. 452 (1949).
(2) Rouse, H. *Fluid Mechanics for Hydraulic Engineers*, Chapter 1, Section 5 (New York: McGraw-Hill Book Co. Inc., 1938).
(3) Hinze, J. O., and Van der Hegge Zijnen, B. G. *Appl. Sci. Res.*, **A1**, p. 435 (1949).

The sources of electron-induced contamination in kinetic vacuum systems

By A. E. Ennos, M.Sc., A.Inst.P., Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks.

[*Paper first received* 1 *July, and in final form* 6 *August*, 1953]

The relative importance as contaminants of various materials which are used in the construction of kinetic vacuum systems was assessed by introducing them into a clean system and measuring the carbonaceous deposit formed on a surface bombarded by electrons. In order of descending contaminating power are diffusion pump oil (hydrocarbon), vacuum grease, various types of rubber gasket material, silicone pump oil, and vacuum wax. Uncleaned metal surfaces are also a considerable source of contamination, and methods of cleaning are discussed. Experiments carried out on the dependence of contamination rate upon the temperature of the bombarded surface showed that a 10 times reduction could be expected from raising the temperature to 200° C. A cold trap at −15° C, surrounding it, was found to be equally effective.

In a previous paper[1] the formation of a contaminating layer on surfaces within a kinetic vacuum system which are struck by electrons has been discussed, with special reference to contamination of electron microscope specimens. It was found that the carbonaceous layer formed was produced by the interaction between the bombarding electron beam and the organic molecules adsorbed upon the surface, carbon being the final product of breakdown of the molecules. These molecules, present on the surface as an adsorbed layer of perhaps only monomolecular thickness at most, are converted to the stable carbonaceous state before a further layer can form on it and itself be broken down by the electron beam. The replenishment of the organic molecule layer was found to be by condensation from the residual vapours within the vacuum system, very little contribution, if any, being made by the migration of molecules along the surfaces surrounding the bombarded areas.

The present paper describes an investigation into the sources from which the contaminating vapours arise. A quantitative estimate of the relative importance of different contaminants has been made, and quantitative information is also given concerning the methods of reducing contamination.

THE VAPOUR SOURCES OF CONTAMINATION

Experimental procedure. The basic requirement for studying the sources of organic vapour which lead to contamination is the provision of a demountable vacuum system which is itself free from sources of contamination, and into which can be introduced samples of possible contaminants. While it is known that glass vacuum apparatus pumped by mercury diffusion pumps and employing a liquid air trap does not exhibit electron-induced contamination when it has been thoroughly outgassed by prolonged baking, it was thought that a less drastic cleaning process might suffice to prevent the formation of carbonaceous deposits. On the other hand, a typical metal demountable system employing rubber gaskets, pumped by a mercury diffusion pump, contaminates as badly as if an oil diffusion pump had been used. The compromise which was made, and which proved effective, was the combination of a glass vacuum system with mercury pump and liquid air trap, and ground glass flange joints sealed with Apiezon W wax on the outer rim. With this construction, the majority of the glassware could be lightly torched while under vacuum, while only a very small surface area of wax of relatively low vapour pressure was exposed to the vacuum. In the initial experiments, a simple three-electrode gun of nickel was arranged to irradiate a polished stainless steel disk held by an additional electrode as in Fig. 1. The anode potential was 2 kV, and the current falling on the target was 100 μA at a density of approximately 15 mA/cm^2. The power dissipation in the target was not sufficient to heat it appreciably and so prevent contamination taking place by the mechanism described in the earlier paper. When contamination occurred, it took the form of a yellow or brown stain on the polished surface of the stainless steel, its configuration and density closely following that of the electron beam where it struck the target.

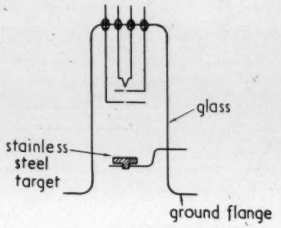


Fig. 1. Simple electron gun with glass envelope

It was found that the apparatus could be rendered clean, so that no contamination deposit was formed, by lightly torching the glass and running a gas discharge inside it. It then remained clean provided that, on demounting it for short periods, air was let in through the cold trap.

On the introduction of such materials as diffusion pump oil, or vacuum tap grease, heavy contamination was formed on the target after running the electron beam for some time (e.g. 100 minutes). On the other hand, the effect of other materials, such as rubber gaskets, metal surfaces, etc., was somewhat variable. It was realized then that the experiments must be carried out under strictly controlled conditions of current density, target temperature, etc., so that the contamination deposit might be assessed quantitatively. It was considered to be a further advantage if more than one "contamination run" could be made without breaking the vacuum.

The apparatus designed to fulfil these conditions is shown in Fig. 2. It consists basically of an electron gun and electrostatic condenser lens focusing the electron beam on to a small slit behind which lies the target, a small piece of glass microscope slide. (The secondary electron ratio of a glass surface at the operating potential of 2 kV is greater than unity, which causes it to assume a positive potential close to that of the anode. Adverse charging-up effects do not then arise.) The target could be moved sideways by means of push rods mounted in side arms and operated magnetically from outside the vacuum. Behind the target was a Faraday cage

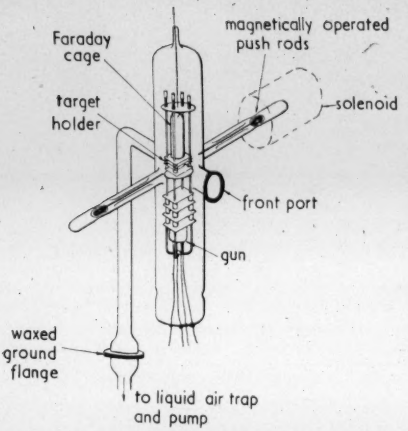


Fig. 2. Apparatus for determination of quantity of contamination deposit

collector for measuring the electron beam. Fig. 3 shows the electrode system diagrammatically. The electrodes are of nickel sheet mounted on glass rod supports with ceramic spacers. The beam-limiting slit consists of two parallel razor blade edges sprung on to the glass target, which has a hole drilled through it to allow the beam to pass through and be measured, when in the central position. The slit limits the beam such that a small patch of contamination 1·4 mm × 0·2 mm is formed on the target.

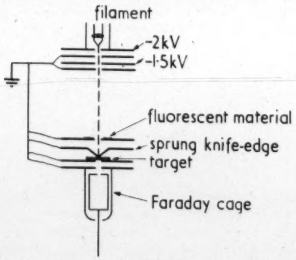


Fig. 3. Electrode system of the contamination apparatus

The whole electrode assembly was enclosed in a glass envelope using mica spacers, and the glass apparatus was mounted within a metal framework so that the heavy magnetic coil used to move the push rods could be slid along support rails. Ground flange joints sealed with Apiezon W wax were employed, as before, for joining the apparatus to the mercury pump and liquid air trap. A similar joint was used to seal on a glass window over the part through which the target and other material was inserted.

The electron gun was capable of giving a focused spot of uniform intensity of 10 mA/cm^2 current density over the beam-limiting slit. The total current reaching the target was 20 μA. This arrangement gave a high current density (and thus a potentially high contamination rate) on the target with small overall heating effect. The target temperature remained low, and thus did not affect the contamination rate.

In operation, the electron beam was first set to the required value with the target in its central position. The target was then displaced by the push rods for the actual run, and moved back again to check the current at the end of the experiment, or between successive depositions.

The thickness of a deposit was measured by removing the target and depositing silver over its surface by vacuum evaporation. The height of the resultant step in the silver film was evaluated by observing fringes of equal chromatic order according to the technique developed by Tolansky.[2] This method of multiple beam interference was used since it allowed measurement over a very small area of deposit without loss in measurement accuracy. Further, the deposit thickness could be measured directly from the wavelength scale of the spectroscope in which the fringes are observed, and its value could be measured at any point along the length of the deposit. This enabled an average value of thickness over the whole length to be taken.

The heavier deposits (e.g. 1000 Å or more) could be measured with considerable accuracy, but for those of the order of 100 Å or less, measurement of the deposit became difficult in view of its possible unevenness and the natural flaws of the glass target surface. The values of less than 100 Å quoted in the subsequent tables are therefore an approximation based on visual appearance. Quoted values of less than 50 Å refer to deposits which are barely visible, and which may be only a few molecular layers thick.

Experimental results. After a number of successive torchings of the glass apparatus while under vacuum, and the prolonged passing of a gas discharge through it, the rate of formation of contamination on the target within the apparatus became negligibly small. It could be maintained in this clean condition as with the simpler apparatus, if the air let into it prior to removal of the window was first passed through a cold trap.

Substances which are likely to be used in the construction of a typical demountable vacuum system were inserted into the front tube of the contamination apparatus, and the target bombarded with electrons for the standard time of 100 min at 20 μA beam current. The quantities of material inserted all relate approximately to those used in the construction of a typical electron gun chamber, e.g. the surface areas of the rubber samples exposed to the vacuum were equivalent to those exposed in such a vacuum system.

Table 1 shows the relative contaminating properties of a number of vacuum materials obtained in this way.

From this table, it is seen that the major contribution towards contamination in a typical demountable vacuum system comes from the pump-oil vapour, the vacuum grease used, and certain types of rubber employed in the gaskets.

Degreasing of the rubber gasket materials by boiling in potash reduced the contamination they produced in the uncleaned state. The surface area of rubber exposed to the vacuum was found to have only a small effect on the resulting contamination, e.g. an increase of area by five times resulted in an increase of contamination by only one-third. This indicates that even a very small area of exposed rubber surface inside the vacuum system can maintain a vapour pressure of hydrocarbon vapour close to the saturated vapour pressure. Any heating of the rubber immediately increased the rate of contamination.

Table 1. *Relative contamination caused by vacuum materials*

($t = 100$ *min*, $I = 0{\cdot}01$ *A/cm*^{2}, $V = 2$ *kV*)

Material	*Treatment*	*Thickness of contamination deposit* (Å)
Diffusion pump oil (Apiezon B)	As supplied	1700
Silicone diffusion pump oil (Dow-Corning 703)	As supplied	500
Vacuum grease (Apiezon M)	As supplied	1500
Apiezon W wax (cold)	As supplied	<50
Natural rubber sheet (very lightly loaded)*	Boiled in aqueous and alcoholic potash	1100
White round-section compass cord (Dunlop Rubber Co. Ltd.)	Boiled in aqueous and alcoholic potash	700
Square section para-rubber gasket cord (Caton Ltd.)	Boiled in aqueous and alcoholic potash	750
Black neoprene (heavily loaded to give oil resistance†)	Boiled in aqueous and alcoholic potash	<50
O-ring rubber gasket material (W. Edwards and Co. (London) Ltd.)	Boiled in aqueous and alcoholic potash	600
Polythene	Washed in detergent	1300
Perspex sheet	Superficially cleaned	<50
Bakelite sheet	Superficially cleaned	<50
Photographic plate	Superficially cleaned	<50

* Mix H.6325, Guide Bridge Rubber Co.
† Mix 15316, Guide Bridge Rubber Co.

The heavily-loaded black neoprene gasket material is that normally used in this laboratory in rotating gland seals. It is often the practice to lubricate this type of seal with vacuum grease, which would render it a source of contamination. This practice has, however, been shown to be quite unnecessary when using the black neoprene.[3] The black sealing Apiezon W wax was found to be non-contaminating only so long as it remained cold and its surface shielded from electron bombardment. The same is true for the Perspex sheet and Bakelite. Polythene, which has been suggested as a suitable gasket material, is seen to be a bad contaminant, however. Silicone oil causes less than one-third the contamination due to the corresponding hydrocarbon oil, but the contamination product formed by it will be siliceous in nature and liable to lead to serious trouble due to its insulating property, which would cause it to charge up in the electron beam.

The metal of a vacuum system, if uncleaned, contributes to the electron beam contamination. Fig. 4 shows the results of early experiments with the contamination apparatus in which the thickness of deposit is plotted against time for successive runs, between which cleaning operations were carried out. Vacuum heating of the metal electrodes and the passing of electrical discharges through the apparatus were seen to free it gradually from contamination material. To make a quantitative estimate of the contamination due to the metal surfaces alone, strips of uncleaned metals were inserted into the clean contamination apparatus and the resultant deposits measured. Table 2 shows that brass and aluminium strip in the normal uncleaned state are bad contaminants, but that the source of the contamination can be removed by drastic cleaning of the surface (actual chemical removal of a layer of the surface). It was found that ordinary grease removers (e.g. pure ether, or an aqueous detergent solution) were ineffective in eliminating the contamination from the metals.

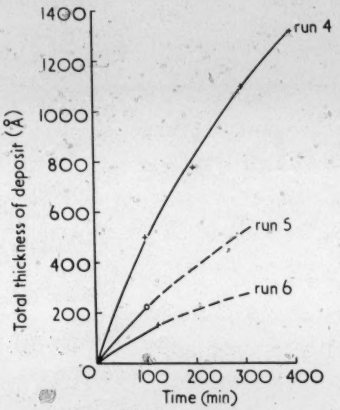


Fig. 4. Results of early experiments with the contamination apparatus

Table 2. *Contamination caused by metal surfaces*

Material	*Treatment*	*Thickness of deposit* (Å)
*Brass strip	Well handled and not subsequently cleaned	700
Brass strip	Cleaned in acid	<50
*Aluminium strip	Well handled and not subsequently cleaned	700
Aluminium strip	Cleaned in aqueous potash	<50
Complete electron gun	Washed in detergent (sodium lauryl sulphate solution)	400
Complete electron gun	Electron beam not operating, cleaned in acid	<100
Complete electron gun	After acid clean, with gun filament running	350
Complete electron gun	As previous expt. but after 6 h baking at 150° C.	<100

* Area of metal equal to internal surface area of electron gun.

By carrying out such drastic surface treatment on the metal surface, it should be possible to construct a complete metal vacuum system which does not contaminate. An electron gun was thus constructed of brass, with glass-metal seals to insulate the leads to the nickel electrodes. Wax joints were used for the other seals, no other organic materials which could cause contamination being employed. The gun was mounted on the clean contamination apparatus to test its contaminating properties. Table 2 also shows that the detergent cleaning was ineffective in reducing contamination, while acid attack of the metal surfaces was completely effective.

confirming the previous results. However, when the filament of the metal electron gun was switched on, further contamination could be detected, indicating that more organic vapour must be driven off from the metal walls, etc., on heating. Prolonged baking of the whole metal gun was then carried out in an attempt to drive off all the vapour. For this purpose, the wax joints and seals of the electron gun were replaced by glass-metal Housekeeper seals, the glass tube connecting the electron gun with the apparatus being waxed on to the apparatus at a considerable distance from the gun, so that the metal could be heated to within the melting point of the solder, while the wax joints remained cool. The result shown in Table 2 can be compared with the measurements of Blears.[4] Blears finds that baking for 4 h at 100° C reduces the partial hydrocarbon vapour pressure by a factor of nearly 10, whereas the contamination rate measurements indicated very little reduction of contamination when baking at this temperature. However, at 150° C the apparatus was rendered almost completely clean after 6 h.

An experiment was carried out to see whether the method of passing a high current discharge inside a metal vacuum system was effective in cleaning up the surface organic material, as it is known to be for glass surfaces. A discharge of 20 mA was passed for 15 min in the brass electron gun, which had no prior special cleaning, and which was attached to the contamination apparatus. Contamination runs before and after the gas discharge showed no significant difference in the amount of the deposit. A metal system thus cannot be rendered free from contamination except by a prolonged vacuum baking of the walls after initial chemical cleaning.

Experiments were conducted to ascertain whether a metal system once cleaned would remain clean after exposure to the atmosphere. The metal system previously described was cleaned by baking and then left open for several days. No appreciable contamination resulted. Any handling or washing did lead to contamination.

The origin of the vapour causing this contamination must be a strongly adherent adsorbed layer on the surface of the metal, since it is seen that only prolonged baking or removal of the top surface layer is effective in eliminating it. Organic vapours from the atmosphere, organic materials (e.g. oils and greases) used in the manufacture of the metal parts, and especially those derived from handling the metal, are presumably the sources of this adsorbed film.

In a survey dealing with the causes of the apparent odour of metals and other minerals, McCartney[5] shows that these odours are caused mainly by adsorbed organic substances, such as mercaptans and thio-ethers which are derived from human perspiration. In handling, or coming into close proximity with the metal, these compounds are transferred to the surface, and can only be eliminated by drastic cleaning (hot acid, nascent hydrogen, or heating to 200–300° C). It appears then that the same organic materials are responsible for the electron beam contamination. At the low pressures within a vacuum system, the adsorbed layers will gradually evaporate off, providing the vapour source of the contamination.

In connexion with the apparent odour of metals, it is interesting to note that the "smell of vacuum" always noticed when readmitting air into a demountable vacuum system is due largely to the metal surface in the vacuum. For example, a piece of uncleaned metal was inserted into the otherwise clean glass system of the contamination apparatus and, on subsequently readmitting air, the strong smell was immediately apparent, even though no vacuum oil, grease or gasket material was present.

REDUCTION OF CONTAMINATION IN NORMAL DEMOUNTABLE SYSTEMS

The experiments described above show that many of the materials used in normal demountable vacuum systems give rise to electron beam contamination, and that only by taking extraordinary care in designing, constructing and cleaning the system can it be avoided. However, as shown qualitatively in the earlier paper, the contamination in a dirty system can also be prevented in either of two ways: (*a*) by heating the surface which is bombarded, (*b*) by surrounding the surface as completely as possible with a cold trap. Further experiments were carried out to find the degree of heating or cooling necessary to provide adequate prevention of contamination in a normal oil pumped vacuum system.

Temperature dependence of contamination. In order to measure the effect of the surface temperature on the amount of contamination formed on the surface, independent of the prevailing vacuum conditions, it was arranged for the electron beam to irradiate a strip of metal having a temperature gradient from one end to the other. The area of bombardment of the metal was confined as before to a narrow strip by means of two parallel knife-edges in contact with the surface, and the focused electron beam was traversed across this slit by deflecting it electrostatically with a saw-tooth potential waveform. The resulting narrow strip of contamination formed was examined optically, and its thickness measured at points along its length.

The arrangement for producing the temperature gradient in the target is shown in Fig. 5. The target is a piece of

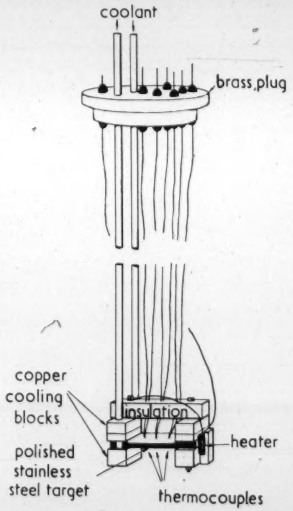


Fig. 5. Arrangement used for production of the temperature gradient in the target

stainless steel strip polished optically flat on a pitch lap, and clamped between an electrical heater at one end and cooled copper blocks at the other. Thermocouples, fitting into a small holes in the target, allow the temperature to be measured at the two ends and in the centre. Cooling liquid can be passed through the support tubes, which are brazed into a brass sealing-plug. This plug fits into the neck of a vacuum

bell-jar, such that the target is facing downwards, inside the vacuum. A 2 kV gun sends the electron beam in a vertical direction onto the target.

The temperature gradient in the target was first verified to be approximately linear when thermal equilibrium had been reached. A number of different temperature ranges could be obtained by adjusting the heater current and coolant temperature (water or acetone-solid carbon dioxide were used). Due to the poor thermal conductivity at the ends of the target, it was not possible to keep the "cold" end below 50° C while maintaining an appreciable temperature gradient. Fig. 6 shows the relationship between the contamination rate and temperature as measured by this method. Between

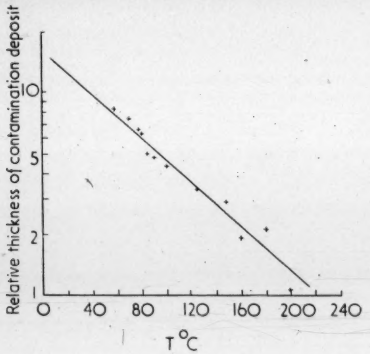


Fig. 6. Temperature dependence of contamination

50 and 200° C, there is a reduction in the contamination rate of approximately 8 times. By extrapolation, the reduction from room temperature (20° C) to 200° C would be about 10 times. This reduction in contamination rate should be of considerable help in high resolution electron microscopy. The use of a hot stage has been attempted by Haine and Mulvey in this Laboratory and proved to give a very marked reduction in contamination.

Prevention of contamination by a specimen cold trap. A brass cold trap surrounding the specimen in an electron diffraction camera was constructed, such that cooling fluid could be circulated through it. The arrangement is shown diagrammatically in Fig. 7. Observations on the contamination of zinc oxide smoke crystals were made by giving a prolonged bombardment to the specimen and then examining it in the electron microscope. Whereas heavy contamination

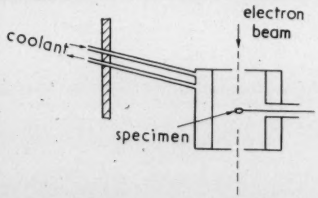


Fig. 7. Brass cold trap enclosing the specimen in an electron diffraction camera

resulted when the trap was not cooled, virtual elimination of it was achieved by circulating a freezing mixture of ice and salt at −15° C. Vapour molecules entering the trap through the apertures are partially condensed on the cooled walls, the resulting drop in partial pressure of hydrocarbon vapour leading to the reduction in contamination. It appears that the relatively high temperature of −15° C of the cold trap is adequate to prevent contamination. Further evidence supporting the effectiveness of a cold trap to prevent contamination by oil was provided by an experiment with the clean glass apparatus in which diffusion pump oil was boiled in a side tube, vapour being trapped by the liquid air trap before it could reach the main chamber. No contamination was obtained.

CONCLUSIONS

The measurements on the contaminating properties of the materials used in the construction of demountable vacuum systems make it possible to specify the conditions under which such a system can be rendered relatively free from the electron beam contamination effect. These are:

1. The provision of a cold trap using coolant of at least −15° C to the oil diffusion pump. The cold trap would have to be perpetually maintained at a low temperature to prevent intake of backing-pump or diffusion-pump oil vapour.

2. The elimination of all high-vacuum grease, the smallest amount of which is a bad source of contamination.

3. The elimination of rubber gaskets except those of the hard black neoprene type, which have been thoroughly degreased. Joints can be made, instead, using soft metal (e.g. solder wire) gaskets, or metal glass seals with waxed glass flange demountable joints.

4. The initial cleaning of all metal parts by superficial removal of the surface (e.g. by acid treatment), followed by pre-baking of those parts of the system which are likely to become warm in operation (e.g. the metal parts surrounding the filament of the electron gun).

While fulfilment of these conditions has been carried out in the case of a small electron gun, rendering it contamination free, in the case of complicated demountable vacuum systems it would be extremely difficult to comply with conditions 3 and 4. In this case, a better approach to the problem of the prevention of contamination is obtained from the two methods discussed in the section on reducing the contamination in normal demountable systems. The surrounding by a cold trap of the specimen and objective aperture (the two most important components to be kept contaminant-free for high resolution microscopy), although difficult in existing microscopes, might be incorporated into future designs.

ACKNOWLEDGEMENTS

The author wishes to express his thanks to Dr. G. Liebmann and Mr. M. E. Haine for many helpful discussions, and to Dr. T. E. Allibone for permission to publish this paper.

REFERENCES

(1) Ennos, A. E. *Brit. J. Appl. Phys.*, **4**, p. 101 (1953).
(2) Tolansky, S. *Multiple Beam Interferometry of Surfaces and Films* (London: Oxford University Press, 1948).
(3) Haine, M. E. *J. Instn Elect. Engrs*, **94**, (Part I), p. 447 (1947).
(4) Blears, J. *J. Sci. Instrum., Supplement No.* 1, *Vacuum Physics*, p. 36 (1951).
(5) McCartney, W. *The International Perfumer*, No. 3, p. 7 (1951).

An improved experimental iteration method for use with resistance-network analogues

By G. LIEBMANN, D.Phil., F.Inst.P., and R. BAILEY, B.Sc., Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks.

[*Paper received* 1 *July*, 1953]

When solving the wave equation $\nabla^2 U + k^2 U = 0$, and similar field problems, on resistance-network analogues by iteration, an error signal can be derived which is a measure of how far the analogue solution satisfies the given equation. An experimental apparatus is described for displaying simultaneously, on a cathode-ray tube screen, the error signals ϵ for a number of resistance-network points; the solution then proceeds until all $\epsilon \simeq 0$. The apparatus can also be used to display network voltages or gradients, which is helpful in many other kinds of problems solved with resistance-network analogues, and to obtain solutions of "best fit" in equation solvers.

It has been shown by one of the authors[1, 2] that the differential equations

$$\nabla^2 U + k^2 U = 0 \tag{1}$$

and

$$\nabla^2 U + \left(k^2 - \frac{n^2}{r^2}\right) U = 0, \; n = 0, 1, 2, \ldots \tag{2}$$

which describe vibration or oscillation problems, e.g. the vibration of membranes or the electromagnetic field oscillations in waveguides and cavity resonators, can be solved by an iteration method in conjunction with a resistance-network analogue. In equations (1) and (2), ∇^2 is the Laplacian operator, $k = 2\pi\nu/v = 2\pi/\lambda$, a characteristic value determined by the frequency ν of the oscillation and its speed v of propagation (or its wavelength λ), and U a physical magnitude (e.g. displacement or magnetic field intensity, etc.) which is represented in the network model by the voltages appearing at the network nodes. In the previously described iteration process, currents I are fed into the resistance-network at the network nodes and the voltage distribution U is measured in the network. From this, an improved value for the "characteristic value" k is found by slide rule calculation. The currents are then re-adjusted, the voltages re-measured, etc., until a "self-consistent" distribution of fed-in currents and resulting network voltages is established, such that at every point the relations

$$I = h^2 k^2 U / R_N \tag{3}$$

or

$$I = h^2 (k^2 - n^2/r^2) U / R_N \tag{4}$$

are satisfied [ref. (2), equation 36], R_N being the local resistance value of the network, h the mesh interval, and r the off-axis distance (in axially symmetric problems).

At the end of the previous paper,[2] it was already mentioned that this iteration process could be greatly accelerated by a different measuring technique. This method will be discussed in this paper, and an apparatus will be described which achieves a great simplification and shortening of the iteration process as used with the network analogue. This apparatus has also been found very useful in certain other classes of problems solved on the resistance-network analogue.

PRINCIPLE OF THE NEW METHOD

Equation (4), which contains equation (3) as a special case ($n = 0$), can be rewritten in the modified form:

$$IR_N - h^2 (k^2 - n^2/r^2) U = \epsilon \tag{5}$$

where the "error voltage" ϵ is a measure of the degree to which equation (4), and hence equation (2), is satisfied. At the beginning of the process of solving equation (1) or (2), $\epsilon \neq 0$ at the various network nodes; the aim of the iteration process is to make $\epsilon \to 0$ simultaneously for each network node as quickly as possible. If the error voltage ϵ is displayed at the same time for all network points, one sees at once at which places in the network the errors are greatest; instead of making cyclic adjustments, one can therefore always work on the worst errors, and reduce these by making appropriate adjustments to the fed-in currents. In this respect, the new method compares with the old network iteration method as the relaxation method[3] compares with the older iterative methods[4] in numerical solutions of field problems. However, it would be an improvement over the relaxation technique in so far as one has under observation *all the time* the visual display of *all the error voltages*, corresponding to the residuals of the relaxation technique, which is of great help in accelerating the process of solution, as it is then always possible to make the *optimum* (temporary) adjustment of the fed-in currents. (It will be remembered that in the iteration process as applied to resistance-network analogues, it is not necessary to re-calculate periodically the characteristic value k from Rayleigh's principle as in the relaxation method proper.)

One can see from the structure of equation (5) that if $\epsilon > 0$, one has to decrease the fed-in currents I or to increase the constant k^2. An increase in I affects mainly the voltage at the network point where the adjustment is made, and to a minor extent the network voltages at its neighbouring points. By contrast, an adjustment of k^2 affects all network voltages, in the same direction. Therefore, if the error voltages ϵ are predominantly of one sign, positive or negative, an adjustment of k^2 is called for, whereas the currents I are adjusted if the sum of the positive and the sum of the negative values of ϵ seem to balance each other approximately. Thus, it is quite easy for the operator to make those adjustments which will most quickly lead to the desired result, $\epsilon \simeq 0$ everywhere.

BASIC RESISTANCE-NETWORK ARRANGEMENT

The basic resistance-network arrangement for carrying out the solution process just described is illustrated by Fig. 1. This shows a "resistance-star," with node P_m representing the mth network point, the local value of the resistances comprising this "star" being R_N, and the network voltage at P_m being U_m. (In an (x, y)-network, R_N has the same value at each network node, whereas in an (r, z)-network, R_N is graded as $1/r$, where r is the off-axis distance of the network point P_m.) The point P_m is connected through a series resistance of value R_N and a rheostat R_x to the voltage source $E_0 (E_0 \gg U_m)$, the current I_m fed into P_m being adjusted by re-setting R_x; preferably $R_x \gg R_N$ to avoid interaction of the current adjustments at different network points P_m. The voltage drop on the series resistance R_N, between points A_m and P_m, is then $E_1 = I_m R_N$; point A_m is connected to the terminal T_1 through the multi-pole switch S_1, ganged with the

multi-pole switches S_0 and S_2. The network point P_m is also connected, in problems described by equation (2), with $n \neq 0$, through a potentiometer resistance R_m, $R_m \gg R_N$, to the zero of the voltage source, the potentiometer tapping point B_m being set to a resistance value, measured from P_m, of $(n^2h^2/r^2)R_m$. Thus, the potential difference $E_2 = (h^2n^2/r^2)U_m$

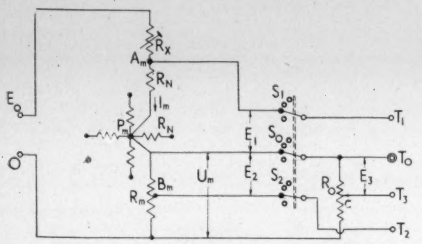


Fig. 1. Basic resistance-network arrangement

$$E_1 = I_m R_N.$$
$$E_2 = \frac{h^2n^2}{r^2} U_m.$$
$$E_3 = k^2h^2U_m.$$

appears between P_m and B_m. The point B_m is joined to the terminal T_2 through a switch S_2. The network point P_m is also connected through the switch S_0 to the terminal T_0, and a potentiometer of value R_0, $R_0 \gg R_N$, is connected between T_0 and the zero of the voltage source. The continuously adjustable tapping point C of this potentiometer is set to a resistance value $k^2h^2R_0$; then the potential difference $E_3 = k^2h^2U_m$ is measured between tapping point C (or terminal T_3) and terminal T_0. A high impedance voltage subtraction circuit, described in the next section, is connected between the terminals T_1, T_2, and T_0, T_3, to perform the operation $(E_1 + E_2) - E_3$. But this expression $E_1 + E_2 - E_3 = \epsilon$, yields the error voltage to be measured. This error voltage ϵ is then amplified and displayed as y-deflexion on a cathode-ray tube with a long persistence screen. The ganged multi-pole switches S_0, S_1, S_2 have as many poles as there are network points P_m, and the error voltages ϵ for all network points P_m are displayed successively; by moving the spot on the cathode-ray tube screen synchronously in the x-direction (e.g. by a further multiple switch ganged to S_1, S_2, S_3), and switching over fast enough, a simultaneous display of all error voltages ϵ according to sign and amplitude can be achieved. The process of solution is then simply to adjust the rheostats R_x to change I_m, or the tapping point on potentiometer R_0 to change k^2 until all $\epsilon \simeq 0$. The potentiometer R_0 can be calibrated to read the constant k directly.

In problems described by equation (2), with $n > 1$, $E_2 > U_m$ for certain network points which would make the discussed potential division by potentiometers impossible as this requires $E_2 \leqslant U_m$. This difficulty can be overcome by inserting an amplifier of gain n^2 between S_2 and terminal T_2, and similarly for the relatively rare case $k^2h^2 > 1$. An easier way of dealing with such cases is to choose a smaller value of resistance αR_N, $\alpha < 1$, between points A_m and P_m. The voltages E_1, E_2, E_3 then correspond to $\alpha I_m R_N$, $(\alpha h^2n^2/r^2)U_m$ and $(\alpha h^2k^2)U_m$, the error voltage becoming $\alpha\epsilon$, and one can always select α small enough when setting up the problem on the resistance-network to make the multiplying factors of U_m everywhere <1.

THE DESIGN OF THE DISPLAY APPARATUS

A block schematic diagram of the complete equipment, as used in practice, is shown in Fig. 2. The terminal T_0 is earthed and becomes the voltage reference point. The voltage measured at T_1 is then $+E_1$ and the voltages at T_2 and T_3 are $-E_2$ and $-E_3$ respectively. The voltages E_1 and $-E_2$ are fed into a first subtracting circuit, which produces an output $-\frac{1}{2}(E_1 + E_2)$. The voltage U_m is passed through a calibrated adjustable potentiometer having a ratio $k^2 : 2$, and the resultant voltage $-\frac{1}{2}E_3$ is fed with the voltage $-\frac{1}{2}(E_1 + E_2)$ into the second subtracting circuit which

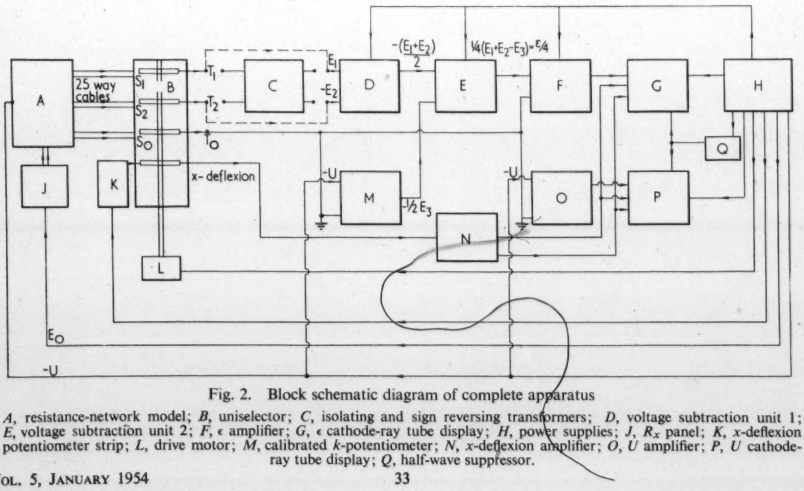


Fig. 2. Block schematic diagram of complete apparatus

A, resistance-network model; *B*, uniselector; *C*, isolating and sign reversing transformers; *D*, voltage subtraction unit 1; *E*, voltage subtraction unit 2; *F*, ϵ amplifier; *G*, ϵ cathode-ray tube display; *H*, power supplies; *J*, R_x panel; *K*, x-deflexion potentiometer strip; *L*, drive motor; *M*, calibrated k-potentiometer; *N*, x-deflexion amplifier; *O*, *U* amplifier; *P*, *U* cathode-ray tube display; *Q*, half-wave suppressor.

produces an output $\frac{1}{4}(E_1 + E_2 - E_3) = \frac{1}{4}\epsilon$. This error voltage $\frac{1}{4}\epsilon$ is amplified and displayed on a cathode-ray tube.

A Post Office type uniselector, which is basically a multipole rotary switch driven by a ratchet motor and having banks of 25 contacts, was found suitable for the present apparatus. By energizing the motor coil with a half-wave rectified 50 c/s supply, it is possible to rotate the wipers twice per second so that the switching rate is 50 positions/second. Under these conditions, the time of transition between switch positions is quite short (about 1 msec). Spurious switching signals of appreciable magnitude but short duration are developed (probably due to changes of contact potential and thermal effects). The appearance of these spurious signals on the cathode-ray tube screen is suppressed by applying a synchronized blackout pulse to the cathode-ray tube modulator for about 4 msec from the beginning of switching.

Details of the two identical voltage subtraction units, following each other as shown in the block diagram, Fig. 2, are given in Fig. 3. If r_{a1}, r_{a2}, μ_1 and μ_2 are the anode impedances and amplification factors of the two similar

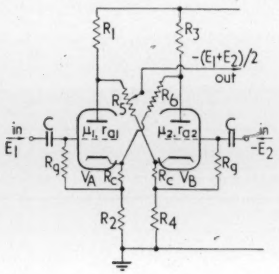


Fig. 3. Schematic diagram of voltage subtraction unit

triodes V_A and V_B, and $R_1 = R_2 = R_3 = R_4 = R \ll R_5 = R_6$, with the cathode resistances $R_c \ll R$, the two valves act as phase splitters having a gain of

$$\left[1 - \frac{2R + r_a}{(\mu + 2)R + r_a}\right] = (1 - \Delta) \qquad (6)$$

Thus, the potential of the anode V_A is $-E_1(1 - \Delta_1)$ and that of the cathode of V_B is $(-E_2)(1 - \Delta_2)$. Hence the potential of the midpoint of R_5 is

$$\tfrac{1}{2}[(-E_2)(1 - \Delta_2) - E_1(1 - \Delta_1)]$$

The effect of the finite values of R_5 and R_6 is to introduce unbalance. The anode load of V_A is R_1 in parallel with R_5 in series with the output impedance of the cathode follower while the cathode load is shunted by R_6 in series with R_3 and the anode impedance of V_B connected as a cathode follower in parallel. This unbalance as well as the unbalance caused by differences in r_a and μ may be taken up by adjustment of the slider of R_5 so that the output is then equal to $-\frac{1}{2}(E_1 + E_2)(1 - \Delta_0)$, where Δ_0 is of the order of

$$\frac{2R + r_a}{(\mu + 2)R + r_a}$$

With the valves and loads in use ($V_A = 6L19$, $R_c = 470\,\Omega$, $R_1 = 47\,k\Omega$, $R_5 = 220\,k\Omega$) the value of $\Delta_0 \simeq 0{\cdot}04$. The calibration of the k-potentiometer takes into account the factor $(1 - \Delta_0)$ and thus makes the instrument direct reading.

The output of the second voltage subtraction unit passes through a variable attenuator (ϵ-amplifier gain control) to the main amplifier. This consists of a low noise pentode directly coupled to a double triode balanced "see-saw" circuit which feeds the Y-plates of the ϵ-display cathode-ray tube.

The x-deflexion voltages are derived from a linear potentiometer strip across part of the h.t. supply, the 25 tapping points of the potentiometer strip being connected to the fourth bank of the uniselector. The potentiometer voltages are passed through a push-pull amplifier and are then applied to the x-plates of both display cathode-ray tubes.

The blackout pulses for suppressing the spurious switching signals are obtained by passing the voltage from the transformer supplying the uniselector motor through a delay network (to compensate for the inertia of the switch), amplifying and clipping it by a double triode valve, and applying it to the grids of the two cathode-ray tubes; diodes across the grid leak resistors provide d.c. restoration.

It is often convenient to use 50 c/s excitation of the resistance-network. The display of the error voltages ϵ would then consist of a series of vertical lines of various length extending symmetrically on either side of the mean level, without an indication of the sign of ϵ. The blackout pulse used to suppress the spurious switching signals is, therefore, extended to suppress a complete half wave of the signal. The extended pulse then removes either the upper or lower half of each line depending on the phase of the signal, i.e. on the sign of the error voltage ϵ. Thus, the display gives direct information about both sign and magnitude of the residual errors ϵ.

In certain applications, d.c. excitation of the resistance-network is more suitable. In this case, the rotating switch acts as a "chopper," and the R-C coupled voltage-subtraction circuits and amplifiers can be used unchanged, an appropriate time constant of 1 sec being used throughout the amplifier chain. The display then consists of a number of bright dots, their relative position with reference to a horizontal line, corresponding to $\epsilon = 0$, giving the value and sign of ϵ. This level $\epsilon = 0$ is marked on the cathode-ray tube screen, and is easily checked by temporarily short-circuiting terminals T_0, T_1 and T_2. Similarly, it is useful to rule vertical index lines on the cathode-ray tube screen corresponding to the switch positions 1, 2, . . . 25; it is then very easy to identify the R_x-rheostat which has to be adjusted to modify ϵ at a particular display position. A typical display at the beginning of a solution for a.c. and for d.c. excitations (otherwise identical conditions) is shown in Figs. 4(*a*) and 4(*b*), the value of k^2 set on the potentiometer R_0 being approximately correct. If the value of k^2 were too small, for otherwise unchanged conditions, the display would look as shown in Fig. 4(*c*) (for a.c. excitation). When the solution has been completed, the display consists of a row of bright dots along the line $\epsilon = 0$, the lines for a.c. excitation having contracted, or for d.c. excitation the dots having all moved towards this line. A disadvantage of the display with d.c. excitation is that the signals pass through the amplifiers only when the rotating switch is in operation, whereas it is possible with a.c. excitation to stop the rotating switch and look for any desired length of time at a particular error voltage ϵ_m.

As is indicated in the block diagram, Fig. 2, the network voltages U also can be displayed, on a second cathode-ray tube. In the block diagram are also shown isolating and sign-reversing transformers, which are by-passed in the applications discussed so far. These isolating high impedance transformers are required in certain network problems in which the voltages E_1 and E_2, which are to be subtracted, do

not possess a common point P_m. This may arise, e.g., in problems involving cascaded networks.[5]

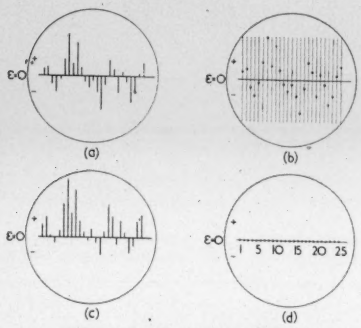


Fig. 4. Display of error voltages ϵ

(*a*) At beginning of solution; a.c. excitation; k^2 approximately correct. (*b*) At beginning of solution; d.c. excitation; k^2 approximately correct. (*c*) At beginning of solution; a.c. excitation; k^2 too small. (*d*) When solution is complete; a.c. or d.c. excitation.

The practical form of the described apparatus comprises four units mounted on standard 19 × 10½ in. Post Office type sub-panels which are assembled in a standard Post Office type rack. The lowest panel carries the power supplies, the second one the rotating switch, sign-reversing transformers, difference units and calibrated *k*-potentiometer, and the two top panels the ϵ- and *U*-amplifiers and display cathode-ray tubes, one on each panel. Access to the 25 ways of each bank of the multiple rotating switch is through a screened multi-core cable terminating in a multiple connector socket, or alternatively in a terminal strip with soldering lugs to which the resistance-network junctions can be connected.

APPLICATIONS OF THE DISPLAY APPARATUS

The primary purpose of this apparatus was the simplification and shortening of solutions of waveguide and cavity resonator problems. It was found that the saving of time and effort is very considerable. For instance, the solution of a relatively simple problem, described by the wave equation, equation (1), with 25 current feeding points (= internal mesh points), took about 3 hours by the previously described network iteration method, after the model had been set up. Using the new display apparatus, this time was reduced to about ½ hour; the time actually spent on establishing the solution within the network to better than 0·2% is only a third of this time, about 10 minutes. The rest of the time mentioned is needed for recording the network results for the 25 internal points (10 minutes) and for applying the correction to the experimental value of k, to eliminate the "mesh size effect," as discussed in ref. (2).

There are many other types of problem in which an apparatus of this kind is very useful. For instance, the task often arises of finding boundary conditions for a problem which produce a given specified field distribution. To give a concrete example from electron optics, one may wish to find the electrode shapes or the electrode potentials of an electrostatic electron lens which give a desired potential distribution along the axis of the system. If it should not be possible to achieve the prescribed axial potential distribution within sufficient accuracy, one may then want to answer the question: how closely can one approach the desired distribution, using a given limited number of lens elements? In a problem of this kind, one sets up the desired potential distribution on a "dummy" potentiometer strip ($= -E_2$) and uses the apparatus to display the error voltage $\epsilon = E_1 - E_2$, i.e. the difference between the voltage in the network, E_1, and the desired voltage value, E_2. This requires only two of the three synchronized rotating switches S_0, S_1, S_2, and the third switch can therefore be used to display (*via* the *U*-amplifier) on the *U*-display cathode-ray tube any other voltage distribution, e.g. the voltage distribution along the boundary points. Thus, one can quickly ascertain the correlation between changes in boundary values and changes in the field distribution at specified interior points.

Another important class of problems where the new experimental apparatus is very helpful is that class where boundary values and boundary gradients are simultaneously prescribed. This arises, for instance, in the resistance-network solution of stress problems.[5] In such a case, one sets up one of the boundary conditions on the network, and displays the difference ϵ of the other specified boundary condition as prescribed (and as set up on a "dummy"), and as actually measured in the network. One can then manipulate the network conditions in the optimum way, watching the display of the error voltages on the cathode-ray tube screen, to establish the desired solution with the least amount of effort.

Obviously, the display apparatus could also be used to simplify and speed-up the work of solving systems of linear algebraic equations with the help of adjustable resistance-networks[6,7] as only one instrument displaying the error voltages ("residuals") simultaneously has to be watched; this would be particularly advantageous in cases where the equations are not "well conditioned," and it is necessary to proceed in small and carefully planned steps of adjustment of the potentiometer settings representing the solution of the problem. Another useful application could be to cases where more equations than unknowns are given and the solution of "best fit" is sought, often required in the interpretation of the results of several series of measurements.

A typical example of this is mass spectrometric analysis, where one uses, in order to reduce the effect of the experimental errors, a greater number of peak heights at the various mass numbers (= number of equations) than the number of constituents (= number of unknowns), and one seeks the "best" solution for the relative abundance of the constituents.

ACKNOWLEDGEMENTS

The authors are indebted to Messrs. A. C. Heath and D. Matthews for help with the construction and testing of the apparatus, and Mr. M. E. Haine for helpful discussions. They also wish to thank Dr. T. E. Allibone, the Director of the Laboratory, for permission to publish this paper.

REFERENCES

(1) Liebmann, G. *Nature* [*London*], **164**, p. 149 (1949).
(2) Liebmann, G. *Proc. Instn Elec. Engrs*, **99**, Part IV, p. 260 (1952).
(3) Southwell, R. V. *Relaxation Methods in Engineering Science* (Oxford: Clarendon Press, 1940).
(4) Liebmann, H. *Sitzber. Bayer. Akad. München*, p. 385 (1918).
(5) Liebmann, G. *Nature* [*London*], **172**, p. 78 (1953).
(6) Walker, R. M. *Proc. Inst. Radio Engrs*, **37**, p. 1467 (1949).
(7) Waldmann, L. *Z. Naturforsch.*, **4a**, p. 226 (1949).

The high-pressure glow discharge in air

By W. A. Gambling, B.Sc., and H. Edels, B.Sc.Tech., Ph.D., A.M.I.E.E., Dept. of Electrical Engineering, University of Liverpool

[*Paper received* 12 *June*, 1953]

The paper describes observations on the high-pressure glow in air between copper and tungsten electrodes and gives characteristics for discharge lengths from 0 to 8 mm and currents from 0·01 to 0·5 A. Measurements of the current density at the cathode and in the positive column are given and a method of estimating the cathode-fall voltage is discussed.

The glow discharge is normally obtained at low gas pressures, but it is known[1,2] that it can be maintained in a similar form if the pressure is increased up to and above 1 atm. The discharge in this pressure range ($\simeq$ 760 mm of mercury) is usually called a "high-pressure" glow discharge. The experimental data so far available[1-5] shows that both low- and high-pressure glows have a high cathode-fall voltage and low cathode-current density compared with those which normally occur in arcs. The high-pressure glow is essentially a low current discharge and an increase in current eventually produces either a sudden or a continuous transition to an arc state characterized by an increase in the current density and a considerable fall in the discharge voltage, which is associated with a changed cathode mechanism of electron emission. Glow-to-arc transitions can also be obtained when separating current-carrying contacts,[6] and in a.c. discharges.[7]

Very few results are available which give a complete set of voltage-current-length characteristics for a high-pressure glow. Such a set of characteristics has been obtained by Wehrli[3] for tungsten electrodes in pure nitrogen. However, since many discharge devices operate in air, the properties of the high-pressure glow in this gas are of considerable interest, and apart from a few isolated measurements the only characteristics available are those given by Thoma and Heer.[1] Unfortunately they did not extend their results either to small currents (less than 0·15 A) or to short discharge lengths (less than 4 mm), and it is shown here that their conclusions regarding the cathode-fall voltage and the uniformity of the voltage gradient along the positive column are incorrect.

APPARATUS

Due to the high cathode-fall voltage and the high voltage gradient in the column at low currents, the discharge can only be maintained by a high voltage supply. The source used was a 1000 V d.c. generator, the output of which was smoothed giving a ripple less than 0·1%. In order to avoid a large increase in current due to the drop in discharge voltage when a transition to an arc occurs, a constant current device was connected in series with the discharge. This consisted of a bank of diode valves connected in parallel and operated at saturation current. The discharge current was then controlled by variation of the filament current of the valves.

The electrodes were supported in a variable gap electrode holder in free air. The gap length was measured by casting a magnified image of the electrodes on a calibrated screen and the magnification of the system was found by substituting a micrometer for the electrodes. In this way it was possible to measure the gap to $\pm$0·05 mm. Two grades of copper were used as electrodes, ordinary commercial copper rod and spectroscopically pure copper, both 7 mm in diameter. The electrodes were cleaned and faced in a lathe and the surfaces polished with fine grade emery paper. Microscopic examination showed that no particles of emery remained embedded in the surface of the metal. No difference in the results was observed for the two types of copper. The tungsten electrodes were 12 mm in diameter; the surfaces were ground flat and polished with emery paper. The discharge was initiated by touching the electrodes together, setting the current to the required value, and then drawing the electrodes apart. Readings were taken keeping the current constant and noting the discharge voltage as the electrode separation was varied. The discharge voltage was measured by a voltmeter connected directly across the discharge.

With copper electrodes several repeatable and consistent measurements could be obtained before cleaning and re-polishing the electrode surfaces. With tungsten, however, oxidation occurred rapidly at currents greater than 0·1 A and it became necessary to clean the electrodes before each reading.

In attempting to extend the characteristics of the discharge into the arc state it was found that at currents less than 1 A the electrodes had to be heavily oxidized before even a comparatively stable arc could be obtained. Even so the arc moved rapidly over the electrode surface and often changed to a glow for a short time causing the voltage to fluctuate. The minimum stable value of this voltage was taken as the arc voltage. The voltages for the glow are accurate to $\pm$1% and those for the arc to $\pm$5%.

RESULTS

A photograph of a high-pressure glow in air at 0·1 A between copper electrodes 1·2 mm apart is shown in Fig. 1(*a*). The cathode spot is covered with a blue glow above which is a dark space; these regions correspond to the negative glow and the Faraday dark space of the low pressure discharge. (The cathode dark space at high pressures is too thin to be seen.) On the anode side of the Faraday dark space is the pink positive column which terminates at the anode in a small bright spot. Surrounding the column and the dark space is a diffuse glow which is not easily visible. When a glow is initiated with clean tungsten electrodes, it has the same appearance as with copper electrodes. At currents

(*a*) (*b*)

Fig. 1. The high-pressure glow in air between (*a*) copper, and (*b*) tungsten, electrodes

greater than 0·1 A, however, oxidation occurs, and over a period of a minute or so the voltage rises by some 20 to 30 V, while at the same time a white, intensely bright cone forms at the cathode and slowly extends into the positive column. A photograph of the glow with tungsten electrodes when this cone has formed is shown in Fig. 1(*b*). Due to the continual fluctuation in voltage it was not possible to obtain consistent readings with the cone present.

Copper electrodes

Voltage characteristics. The voltage-length characteristics for various currents and the derived voltage-current curves obtained with copper electrodes are shown in Figs. 2 and 3.

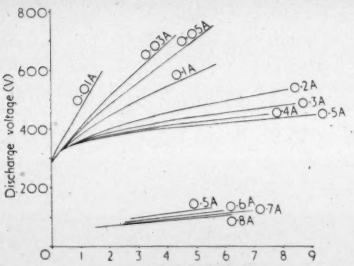


Fig. 2. Voltage-length characteristics with copper electrodes for the glow (upper curves) and the arc (lower curves)

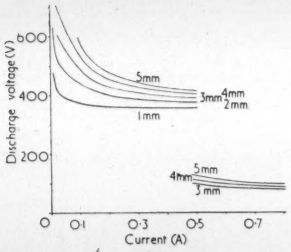


Fig. 3. Voltage-current characteristics with copper electrodes for the glow (upper curves) and the arc (lower curves)

By noting the lowest voltage reading as the electrodes were brought together until they touched, an approximation to the discharge voltage at zero gap length was obtained. The value of this voltage is 285 V and is the same over the range of currents used. For discharge lengths between 0 and 0·3 mm the curves for currents greater than 0·01 A are coincident (within the accuracy of measurement). Since this length corresponds to the visual length of the Faraday dark space plus that of the negative glow, it appears that the voltage drop across this portion of the discharge is about 40 V and is constant for these currents. The voltage gradient in the positive column, as given by the slope of the voltage-length characteristics, is not constant, but decreases to a constant value as the discharge length is increased. This constant value, which is attained at lengths greater than 4 mm, varies from 1500 V/cm at 0·01 A to 85 V/cm at 0·5 A. At long gaps, and particularly with small currents, the anode and cathode spots tend to move causing voltage fluctuations.

Fig. 3 shows that for a fixed gap length, as the current is increased the discharge voltage falls and tends to become constant, while at small currents the voltage increases sharply. For currents down to 0·03 A this voltage increase is caused solely by the increase in voltage gradient in the positive column, since the sum of the electrode drops and the voltage across the Faraday dark space and negative glow is constant. The curves are not continued for currents greater than 0·5 A, but this value has no significance beyond the fact that at higher currents temporary transitions to an arc occurred too frequently to allow accurate readings to be taken. If the glow current were held at, say, 0·2 A then occasional drops in voltage were observed as the discharge momentarily changed to an arc and then reverted to a glow. As the current was increased the frequency of these "temporary" transitions increased until at currents of 0·6 A or more the discharge was in a continual state of oscillation between a glow and an arc. These and other similar voltage disturbances have been observed by Fan.[4] Thus the transition current has not a unique value, but as the current is increased the percentage of time for which the discharge is an arc increases until at about 0·8 A the arc discharge becomes comparatively stable. At small discharge lengths, and at the currents used in the experiments, the arc was very unsteady and it was not always possible to obtain even approximate readings with arcs less than 2·5 mm in length.

Current density measurements.—With copper electrodes the glow leaves a well-defined cathode spot which is usually circular and consists of a central area of clean metal surrounded by a black ring and further rings of different colours. The diameter of the area of clean copper was measured with a travelling microscope and by assuming the current to flow

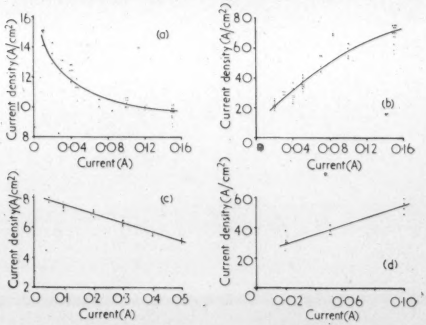


Fig. 4. Current density measurements, with copper electrodes (*a*) at the cathode, (*b*) in the positive column, and with tungsten electrodes (*c*) at the cathode, (*d*) in the positive column

in this region the mean current density at the cathode was determined. The mean current density at the cathode end of the positive column was obtained from measurements of the discharge diameter at this point on the assumption that no current flows in the diffuse sheath which surrounds the positive column. The results of the current density measurements are given in Figs. 4(*a*) and (*b*).

As the current increases, the cathode current density falls from 15 A/cm² at 0·01 A to a nearly constant value of 10 A/cm² at 0·15 A. The results for the positive column are rather scattered, due to the difficulty of judging the edge of the column, but nevertheless they indicate that the current density increases with current. It was not possible to continue these curves to higher currents due to the onset of temporary transitions.

Tungsten electrodes

Voltage characteristics. Figs. 5 and 6 give the voltage characteristics for tungsten electrodes. These are similar to

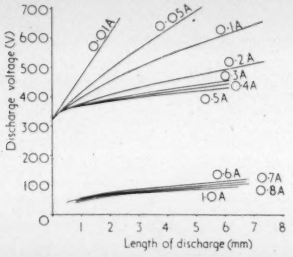


Fig. 5. Voltage-length characteristics with tungsten electrodes for the glow (upper curves) and the arc (lower curves)

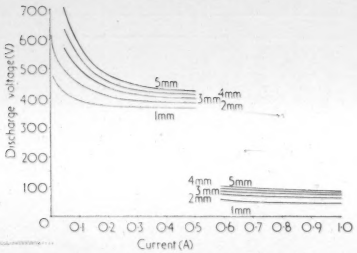


Fig. 6. Voltage-current characteristics with tungsten electrodes for the glow (upper curves) and the arc (lower curves)

those obtained with copper. The limiting glow voltage as the electrodes are brought together is 328 V and is the same at all currents. For discharge lengths between 0 and 0·3 mm the curves for currents greater than 0·01 A are again coincident giving a value for the voltage drop across the negative glow and Faraday dark space of about 25 V, i.e. lower than the value obtained with copper electrodes. The voltage gradient in the positive column decreases as the discharge length is increased and tends to a constant value which is roughly the same as that with copper electrodes at corresponding currents. The arc was more stable with tungsten and it was possible to obtain voltage readings at small arc lengths.

"Temporary" transitions from the glow to the arc also occur with tungsten electrodes, but at a given current the frequency of these transitions is less than for copper. However, because of the rapid formation of oxide on the tungsten it was not possible to obtain consistent measurements at currents greater than 0·5 A, although glow discharges (with the oxide cone referred to earlier) were observed at currents up to 1 A, the maximum current used.

Current density measurements. The central area of the cathode marking on tungsten is discoloured and the boundary between this area and the various surrounding coloured rings is not very clear. However, it was possible to measure the diameter of the cathode spot at currents up to 0·5 A and to measure the mean cathode current density. As shown in Fig. 4(*c*), this falls from about 8 A/cm² at 0·05 A and to 5 A/cm² at 0·5 A. Although clean electrodes were used for each reading, tungsten oxide was forming during the experiment and it is not known how this affects the current density.

The current density in the positive column could not be measured at currents greater than 0·1 A due to the disturbing effect of oxidation, but Fig. 4(*d*) shows that between 0·02 A and 0·1 A it increases from about 30 A/cm² to 55 A/cm². Thus at corresponding currents the current densities in the positive column with copper and tungsten electrodes are approximately the same.

DISCUSSION

The slope of the linear portion of a voltage-length characteristic gives the voltage gradient in the central portion of the positive column, i.e. that part of the column not influenced by the proximity of the electrodes. A comparison of the values so obtained with those found by Thoma and Heer, who studied a glow discharge up to 90 mm in length between water-cooled copper electrodes in air, is given in the table, which shows that the two sets of values for copper electrodes are in good agreement. The values for tungsten are similar to, but not identical with, those for copper. It is interesting to note that the voltage gradient along the arc column is not markedly different from that along the positive column of a high pressure glow carrying the same current, as illustrated by the table.

Voltage gradient along the discharge column in V/cm

	Present results		Thoma and Heer.[1]
Current (A)	*Tungsten*	*Copper*	*Copper*
Glow			
0·01	1450	1540	—
0·05	650	760	—
0·1	350	450	—
0·2	200	185	182
0·3	136	126	130
0·4	125	105	107
0·5	90	86	84
Arc			
0·5	—	≃ 135	
0·6	≃ 90		

At low pressures in the normal glow discharge, i.e. when the cathode is not completely covered, the cathode current density is constant and independent of current. Fig. 4(*a*), which is for uncooled copper electrodes at 1 atm, shows that as the current is increased the cathode current density falls and approaches a constant value of 10 A/cm.² This result may be compared with previous investigations of glows between water-cooled copper electrodes in air. Thus Thoma and Heer[1] did not observe any variation with current, but state that the cathode current density is ≃10 A/cm², while Fan[4] gives a value of 12·5 A/cm² which is independent of current. By measuring the area of the negative glow on the side remote from the cathode, von Engel, Seeliger, and Steenbeck[2] found that the current density was 8·4 A/cm² for currents

greater than 0·1 A and decreased at lower currents. Using the same method we obtained a similar value of 8 A/cm^2 when the current density at the cathode surface was 10 A/cm^2. The difference in the values given by the two methods is due to the slight expansion of the negative glow with distance from the cathode. Thus, although these results show that the cathode current density with copper electrodes is constant for currents greater than 0·1 A, there appears to be no agreement as to the variation at lower currents. It is possible that the difference is due to temperature effects in that the electrodes used by von Engel, Seeliger and Steenbeck were water-cooled, whereas those used in the present experiments were not. With tungsten electrodes the cathode current density does not become constant but continues to decrease as the current is increased up to 0·5 A. However, it is not known how the values obtained are affected by the presence of the oxide cone referred to earlier.

The limiting voltage as the electrodes are brought together is independent of current for both copper and tungsten. This voltage will be referred to subsequently as the "zero-length" voltage. The values obtained are 285 V and 328 V for copper and tungsten respectively, and agree closely with the values 280 V and 325 V obtained by Burstyn.[8] Because they did not study short discharges Thoma and Heer[1] conclude that the voltage-length characteristics for the high-pressure glow between copper electrodes in air are straight lines which intersect on the voltage axis giving a zero-length voltage of 350 V; this they identified as the cathode-fall voltage. Fig. 2 shows that the characteristics are curved for lengths less than 4 mm and hence the zero-length voltage given by Thoma and Heer is incorrect.

The cathode-fall voltage is an important property of the glow discharge and at low pressures it can be determined by means of probe techniques. In this way, with copper electrodes in air, Rottgardt[9] obtained the values 252 V and 295 V depending on whether the probe was situated at the cathode edge of the negative glow or in the Faraday dark space, while by prolonged drying Schaufelberger,[9] with a probe at the anode edge of the negative glow, obtained the much higher value of 375 V. Experiments at low pressures also show that with normal discharges, in the pressure range for which probe techniques are valid, the cathode-fall voltage is constant and the cathode-fall region obeys the similarity laws. At high pressures the use of probes is not possible so that alternative methods of determining the cathode-fall voltage must be sought.

It has been shown by von Engel, Seeliger and Steenbeck,[2] from measurements of the cathode-current density, that the similarity relation holds for the cathode-fall region of the glow between copper electrodes in air from low pressure up to 1 atm. This may be taken to indicate that the cathode-fall voltage remains constant up to high pressure and is given by the value obtained by probes at low pressures. A further indication of the cathode-fall voltage may be obtained from the zero-length voltage, which can be measured at high pressures. If the proximity of the electrodes does not seriously affect the electrode regions then the zero-length voltage is the sum of the electrode falls and approximates to the cathode-fall voltage. For copper electrodes in (moist) air the zero-length voltage of 285 V may be compared with Rottgardt's values of 252 V and 295 V for the cathode-fall at low pressures. The value 295 V is certainly too high since the probe for this measurement was in the Faraday dark space, and the value 252 V obtained with the probe at the edge of the negative glow is probably not accurate since the cathode fall extends some distance into the visible region of the negative glow. Thus the zero-length voltage gives a reasonably accurate estimate of the cathode-fall voltage.

There is some evidence in favour of this view. For instance not only with copper electrodes but also with iron, silver and platinum electrodes in air, the zero-length voltage (Burstyn's measurements)[8] lies between the two values given by Rottgardt[9] for the cathode-fall voltage, which shows that in these cases also, the zero-length voltage gives a reasonable value for the cathode-fall voltage. Furthermore, Schaufelberger[9] has shown that the removal of moisture increases the cathode-fall voltage, and when the zero-length voltage was measured by the authors in dry air an increase of 25 V was obtained. This increase is not as much as that observed by Schaufelberger for the cathode-fall voltage, since with the apparatus available it was not possible to remove completely all the moisture present. Finally, Figs. 2 and 5 show that the zero-length voltage is independent of current and further experiment showed that it also remains constant as the pressure is reduced to 130 mm of mercury.* These two properties are in accordance with the behaviour of the cathode-fall voltage at low pressures.

The evidence indicates therefore that the zero-length voltage may be used as a measure of the cathode-fall voltage. The value so obtained includes the anode-fall voltage, but this is probably small (at low pressures it is $\simeq$ 15 V or zero)[10] and may be neglected. There does not appear to be a published value for the cathode-fall voltage with tungsten electrodes in air, but from a consideration of the zero-length voltage it would appear to be $\simeq$ 328 V.

Although the zero-length voltage is higher in dry air, measurements over a period of five months showed that the discharge was not affected by day-to-day variations in humidity, so that the results of Figs. 2 to 6 give the properties of the high-pressure glow in air under the conditions normally met with in practice.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the many helpful discussions they have had with Professor J. M. Meek and Dr. J. D. Craggs and for their advice in the preparation of this paper. One of us (W. A. G.) is grateful for the financial support provided by the Department of Scientific and Industrial Research.

REFERENCES

(1) THOMA, H., and HEER, L. *Z. Techn. Phys.*, **13**, p. 464 (1932).
(2) VON ENGEL, A., SEELIGER, R., and STEENBECK, M. *Z. Phys.*, **85**, p. 144 (1933).
(3) WEHRLI, M. *Z. Phys.*, **44**, p. 301 (1927).
(4) FAN, H. Y. *Phys. Rev.*, **55**, p. 769 (1939).
(5) BRUCE, C. E. R. *Nature* [*London*], **161**, p. 521 (1948).
SUITS, C. G. *J. Appl. Phys.*, **10**, p. 648 (1939).
(6) HOLM, R. *Electric Contacts*, p. 255 (Sweden: Hugo Gebers Forlag, 1946).
(7) MACKEOWN, S. S. *Elect. Engng*, **51**, p. 386 (1932).
(8) BURSTYN, W. *Elektrische Kontakte*, p. 17 (Berlin: Julius Springer, 1937).
(9) See BÄR, R. *Handbuch der Physik*, Vol. 14, p. 199 (Berlin: Julius Springer, 1927).
(10) DRUYVESTEYN, M. J., and PENNING, F. M. *Rev. Mod. Phys.*, **12**, pp. 170–73 (1940).

* At lower pressures, as the electrodes are brought together the discharge voltage at first decreases and then increases as the anode enters the cathode-fall region. The minimum voltage, however, remains at the value given by the zero-length voltage (285 V).

Approximate electrode shapes for a cylindrical electron beam

By E. R. Harrison, Atomic Energy Research Establishment, Harwell, Berkshire

[*Paper first received* 6 *March, and in final form* 31 *July*, 1953]

For a parallel cylindrically symmetrical electron beam, in which the current is space charge limited, the accelerating electrodes have varying shapes which previously have been determined with the electrolytic tank. With the aid of some approximations a simple equation is obtained for the shapes of such electrodes. The errors involved are a maximum for the cathode surface; for most other electrodes it is found that this equation gives results indistinguishable from those determined experimentally, and can therefore be used in the design of those electrodes for which there are no data.

In a space charge limited current between infinite parallel planes the potential varies as

$$V_0 = Az^{4/3} \tag{1}$$

from the Child–Langmuir equation, where the co-ordinate z is measured from the emitting plane in the direction of the electron flow. If the parallel electron flow is to have a finite cross-section of a particular shape, then in addition to equation (1) there is the condition that ∇V is zero perpendicular to the surface boundary of the electron flow. Only in the case of a plane symmetric flow is there an analytic solution of $\nabla^2 V = 0$, which satisfies exactly the boundary conditions of the beam. The shapes of the cathode and anode electrodes conform to known relations, and these shapes are independent of the distance of the beam edge from the plane of symmetry.

For an electron beam having cylindrical symmetry about the z axis, there is no known analytic solution of the potential field outside the beam, and it is usual to resort to experimental methods for determining the electrode shapes. For a beam of radius r_0 and current I

$$A = [9I/(2e/m)^{1/2}r_0^2]^{2/3}$$

and in addition, the potential must satisfy the conditions

$$r \leqslant r_0\text{:}\quad \frac{\partial^2 V}{\partial z^2} = \frac{4V}{9z^2},\ \frac{\partial V}{\partial r} = 0,\ V(r,z) = V_0 \tag{2a}$$

$$r \geqslant r_0\text{:}\quad \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{\partial^2 V}{\partial z^2} = 0 \tag{2b}$$

With the electrolytic tank Pierce* has plotted the shape of a number of electrodes which fulfil the boundary conditions of $V(r_0, z) = V_0$, $\partial V(r, z)/\partial r = 0$ at the beam edge.

A certain amount of difficulty is experienced, however, when attempting to use these graphical results for the purposes of constructing the electrodes. The cathode has approximately a conical shape and is the easiest electrode to construct accurately, and the situation would be considerably simplified if some elementary equation were available which specified the anode shape and involved relatively negligible errors. Such an equation might eliminate to some extent the necessity for experimentally determining the shapes of anodes for which there are no published data. It is shown below that an approximate equation of this kind may be readily derived.

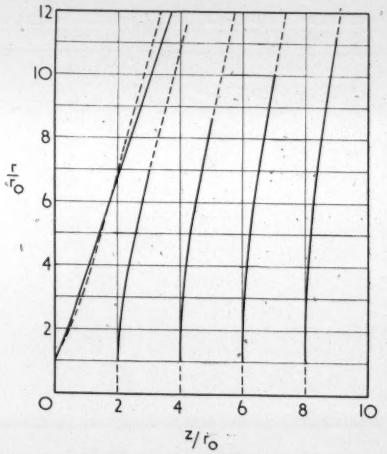
In a solenoidal field of rotational symmetry, let the potential be given by the series

$$V(r, z) = \sum_{n=0}^{\infty} (r - r_0)^n a_n(z)$$

Using equation (2*b*), this series can be expressed in terms of $a_0(z)$ and its differential coefficients, and by putting $V_0 = a_0(z)$, we have

$$V(r, z) = V_0\left[1 - \frac{1(r - r_0)^2}{9z^2(1 - r_0/2r)} + \frac{5}{648z^4}\frac{(r - r_0)^4}{(1 - r_0/2r)(1 - r_0/4r)} - \ldots\right] \tag{3}$$

In order to simplify the calculations we assume that $r - r_0 < z$, and neglect all terms including $(r - r_0)^4/z^4$ and the higher powers. This procedure leaves unaltered the conditions of $V(r_0, z) = V_0$ and $\partial V(r, z)/\partial r = 0$ which must be satisfied at the cylindrical boundary of the beam.



Equipotential surfaces for a cylindrical electron beam of radius r_0. The surfaces are for $z = 0$, $2r_0$, $4r_0$, $6r_0$ and $8r_0$, corresponding to potentials of 0, $0{\cdot}158V_0$, $0{\cdot}397V_0$, $0{\cdot}681V_0$ and V_0, respectively. The full curves are from equation (4) and the pecked curves are Pierce's experimental results

On expanding equation (3) for $z + \Delta z$ by Taylor's series

$$V(r, z + \Delta z) = V(r, z) + \Delta z\frac{\partial V(r, z)}{\partial z} + \frac{\Delta z^2}{2!}\frac{\partial^2 V(r, z)}{\partial z^2} + \ldots$$

* Pierce, J. R. *J. Appl. Phys.*, **11**, p. 548 (1940).

or, neglecting Δz^2 and the successive terms

$$V(r, z+\Delta z) = V_0\left[1 + \frac{4\Delta z}{3z} - \frac{1}{9z^2}\frac{(r-r_0)^2}{(1-r_0/2r)}\left(1 - \frac{2\Delta z}{3z}\right)\right].$$

If now $\Delta z = 0$ at $r = r_0$, the surface of constant potential $V(r, z+\Delta z) = V_0$ cuts the beam edge at z, and conforms to the equation

$$\frac{r-r_0}{(1-r_0/2r)^{1/2}} = \pm 6z\left(\frac{\Delta z}{3z - 2\Delta z}\right)^{\frac{1}{2}} \qquad (4)$$

The error involved is a maximum for the cathode surface where $(r - r_0)/z$ has the largest values. On putting $V(r, z) = 0$ in equation (3), a simple relation is obtained for the cathode surface, from which $dr/dz = 3/\sqrt{2}$ at the beam edge and $dr/dz = 3$ for large values of r, corresponding to angles of 65° and 71·5°, respectively. The "Pierce cathode" has initially an angle of 67·5°, and thereafter increases to approximately 75°. For most values of $V(r, z) > 0$, however, the error in equation (4) is extremely small as can be seen in the figure. This equation is used for plotting r/r_0 against z/r_0 (full curves) for the equipotential surfaces cutting the beam at distances of $z = 0$, $2r_0$, $4r_0$, $6r_0$, and $8r_0$. The pecked curves are reproduced from Pierce's experimental results. As z/r_0 increases, the two sets of curves rapidly tend to become indistinguishable, and equation (4) is sufficiently accurate in practice for most anode shapes.

A proton beam accelerator has been considered, having a series of electrodes shaped for space charge limited flow. The proton beam is extracted from the discharge in the ion source with the correctly shaped electrodes, and for large values of z/r_0 the successive electrodes have a radius of curvature given with sufficient accuracy by

$$R = 2V_0'/V'' = 6z \qquad (5)$$

The preparation and use of cells for the realization of the triple point of water

By C. R. Barber, B.Sc., F.Inst.P., National Physical Laboratory, Teddington, Middlesex; R. Handley and E. F. G. Herington, Ph.D., A.R.C.S., A.R.I.C., Chemical Research Laboratory, Teddington, Middlesex

[*Paper received* 7 *May*, 1953]

The Ninth General Conference of Weights and Measures[1,2] which met in October 1948 passed a resolution that the zero of the International Temperature Scale should be defined as being the temperature 0·0100° C below that of the triple point of water. The technique used at Teddington for preparing cells for the realization of the triple point of water is described and the results obtained by the use of these cells over an extended period are discussed.

Work in these laboratories has confirmed the experience of other investigators that the temperature of the triple point of water can be realized more readily than can the ice point. The triple-point temperature is higher than the ice point and this difference results partly from the effect of pressure and partly from the effect of dissolved gas.

Calculation shows that in passing from the ice point to the triple point the pressure change should result in an elevation of 0·00747° C, while the removal of air having 0·03% of carbon dioxide should produce a rise of 0·00244° C.[3] The triple point of water should therefore be 0·0099° C above the ice point. Four experimenters[4,5,6,7] have reported a value of 0·0098° C and a fifth[8] 0·00997° C. It has been agreed therefore by the General Conference of Weights and Measures[1,2] that the ice point should be taken as being 0·0100° C below the triple point.

Part III, section (*a*) of the text of the International Temperature Scale of 1948[1] states that, "The temperature of equilibrium between ice, liquid water and water vapour has been realized in glass cells from 4 to 7 cm in diameter, which have an axial re-entrant well for the thermometers and contain only water of high purity. The amount of water should be such as to permit adequate immersion of the thermometer and to ensure the existence of the three phases during measurements. Such cells, when properly prepared for use and kept entirely immersed in an ordinary ice bath, have been found to be capable of maintaining a temperature constant to 0·0001 degree for several days." No additional information on making or filling the cells is given in the text.

Ideally the water used in such cells should be absolutely pure, but as it is impossible to obtain such a material it is necessary to consider the minimum purity requirements that the water sample must satisfy. The water should be as free as possible from adventitious ions and other water-soluble impurities. Since the freezing point depression constant of water is 1·86° C per mole of impurity in 1000 g of water[9] it follows that the concentration of soluble impurity should be less than 0·000054 mole per 1000 g of water if the freezing point depression produced by the impurity is to be less than 0·0001° C. This permissible concentration is extremely small and it is obvious that the solubility of the material from which the containing vessel is made must be low. The technique described later for producing the ice-water interface adjacent to the thermometer well does to some extent reduce the effect of small amounts of impurity but obviously the presence of any impurity is a source of uncertainty in realizing the triple point. The cells described here were made of Pyrex and although this glass undoubtedly dissolves in water to some extent[10] no measurable depression of the triple point has been detected on keeping the cells for five years. The presence of grease should be avoided because not only will it depress the triple point insofar as it is soluble, but it may tend to collect on the surface of the ice crystals and modify their growth.

The isotopic composition of the water is of some importance and, indeed, if the triple point is to be employed as an exact temperature standard, the isotopic composition of the water should be completely defined. The elevation in freezing point, Δt, produced by a small mole fraction, x, of deuterium oxide in H_2O is given by the equation, $\Delta t = 4{\cdot}213x$.[11] Since x is approximately 0·00015 for natural waters (see Kirshenbaum)[12] it follows that the triple point of natural

water will be 0·0006° C higher than if there had been no deuterium present. Variations as large as 0·000014 mole fraction in the deuterium content of natural water have been recorded[12] so that the source of water used in the triple-point cells must be carefully selected if the temperature differences between cells are to be less than 0·00006° C, i.e. correct to the nearest 0·0001° C.

Fortunately, however, the deuterium content of tap water in both America and Great Britain[12] is reported to be the same within a few millionths of a mole fraction, so that tap water after purification was used in the present cells. Further, Emeléus and others[13] have shown that the density of water prepared from London Metropolitan Water Board tap water as delivered at South Kensington, London, did not vary by more than ± 0·2 parts per million over a period of six months.

The distillation procedures used for the purification of the water must be designed to minimize fractionation because the various isotopic forms of water have different volatilities. For example, the fractionating factor for the distillation of water at 100° C is reported[12] as 1·03 and hence it may be calculated that a fractionating column equivalent to ten theoretical plates under total reflux will bring about a 33% depletion in the deuterium content of the water at the top of the column, i.e. the temperature of the triple point of the water at the top of the column would differ by 0·0002°C from the original still charge. In view of these considerations the technique of preparing water developed by Emeléus and others,[13] which involved minimum fractionation formed the basis of the method used in the present work.

The cells were cleaned before filling by a procedure widely adopted for preparing electrical conductivity cells, i.e. by treating the cell with chemical cleaning mixture, followed by copious washing with tap water and prolonged steaming. Although the behaviour of the cleaning mixture used, chromic-sulphuric acid, has been adversely critized (Partington)[14] on the grounds that it etches the glass which cannot afterwards be washed free of chromic acid, it was thought that a certain amount of attack of the glass was not necessarily disadvantageous if some of the more readily soluble portions of the glass were removed. Moreover, in view of the observations of Laug[15] it appeared that the prolonged steaming would remove the chromium.

PREPARATION OF WATER

The two stills of capacity 2 litres shown in Fig. 1 were constructed entirely of Pyrex. Each vertical section was 100 cm long and 2·5 cm bore and was packed with short lengths of Pyrex tubing 1 to 1·5 cm long, 0·4 cm internal and 0·6 cm external diameter which were inserted to catch spray. A Pyrex tube sealed at the upper end and also approximately 0·5 cm from the bottom was placed in each flask to reduce bumping. The ground glass joints were employed ungreased; each still after fabrication was washed well with chromic-sulphuric acid cleaning mixture and with tap water. Still No. 1 was charged with 2 litres of tap water to which were added 0·2 g of Analar potassium permanganate and 0·4 g of pure sodium hydroxide. The condenser was turned so that the liquid flowed back into the column and the contents were refluxed for 8 h. The condenser was then rotated so that water could be distilled into a previously steamed Pyrex glass-stoppered flask. The water was distilled until about 500 ml remained in the boiler; this residue was rejected. The distillate from still No. 1 was used to fill still No. 2. No condenser was employed in this second stage and the liquid was always boiled at such a rate that approximately one quarter of the steam reaching the top of the column condensed in a Pyrex bottle and the remainder escaped. This escaping steam acted as a protective blanket and prevented back diffusion of impurities from the air into the bottle and still.

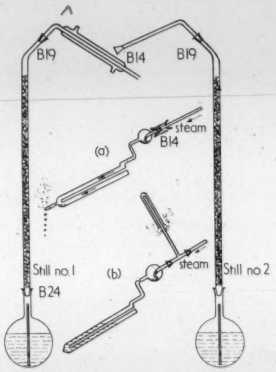


Fig. 1. Preparation of water and technique for filling cells. (*a*) Steaming-out the cell; (*b*) Filling the cell

Slight changes in isotopic composition will occur during the preparation of the water because all the liquid was not distilled in the first stage and because the vapour was only partially condensed in the second stage. Calculation shows that each of these composition changes considered separately can be ignored if the triple point is to be correct to 0·0001° C. Moreover, whereas the method of using the first still produces a slight decrease in HDO content the second still produces a slight increase so that the overall change in isotopic composition is very small. Approximately 10 litres of water was distilled through each newly erected column before the equipment was used to steam and charge triple-point water cells. The water obtained from these two stages of distillation was returned to the empty boiler of still No. 2 and this charge when boiled was used to steam and fill the triple-point cell.

PREPARATION OF TRIPLE-POINT CELL

Construction of cell. Fig. 2 shows two types of triple point cells (*Y* and *Z*) as received from the glass blower. The cells after filling consist only of the sections within the dotted boxes. Cells of type *Y* when completed carry a side arm which forms a convenient handle for clamping. Type *Z* with a projecting thermometer pocket can be more deeply immersed in the ice bath and can be held by means of this pocket. Cells of various dimensions were constructed suitable for thermometers of different sizes. Two typical examples of dimensions were: Type *Y* cell, body of cell 42 cm, external diameter 4·5 cm, thermometer pocket length 37 cm, internal diameter of pocket 1·2 cm; Type *Z* cell, body of cell 28·5 cm, external diameter 4 cm, thermometer pocket length overall 33 cm of which 7 cm projected outside the cell, internal diameter of pocket 0·8 cm. Before cleaning, the cell was tested for the absence of pinholes by means of a Tesla coil.

Cleaning the cell. After closing the port *A* (Fig. 2) with a ground glass cap the cell was filled with a saturated solution of

chromic acid in concentrated sulphuric acid and this solution was allowed to stand in the cell for several minutes before draining. The cell was then washed with tap water, a small quantity being used at first to produce a rise of temperature by the interaction of the water and acid, followed by larger amounts. The cell was finally filled with distilled water and capped at the open end, *B*, until required for steaming. Still No. 2 was charged with doubly-distilled water and the top of the column of this still was provided with a B 14

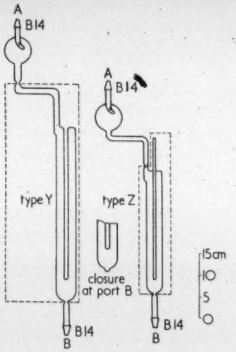


Fig. 2. Triple point cells before filling and sealing

socket so that the triple-point cell could be connected easily in a position where its axis was approximately at an angle of 45° with the horizontal [see Fig. 1(*a*)]. A connexion (ungreased) was made between the still and port *A* (Fig. 2) and steaming was carried out for 8 h when the cell was reversed and the steaming was continued for a further 8 h through port *B* (Fig. 2). Finally port *A* was reconnected and steam was passed for an hour before disconnecting the cell and sealing off at port *B*. This seal was made by collapsing the constricted tube close to the main body of the cell, followed by blowing the seal into the form of a small bulge as shown inset in Fig. 2. During this operation the interior of the cell was protected from contamination by the use of a thin rubber membrane between the cell and the glass blower's mouth. The main object in making this form of seal was to obviate the usual "spiked" seal which has been found more prone to fracture when the cell is immersed in an ice bath.

Filling the cell. Still No. 2 was recharged with doubly-distilled water and the triple-point cell was reconnected, through port *A*, after interposing a tube carrying a side arm to act as a steam vent [see Fig. 1(*b*)]. Part steaming and part condensation were continued over a period of approximately 4 h until the body of the cell was completely filled. The filled cell was then removed and a connexion made at joint *A* (Fig. 2), *via* a glass tap to a rotary oil pump. A very small ring of vacuum grease was placed at the base of the joint at *A*. Suction was now cautiously applied and the cell which was approximately vertical was warmed by steadily waving a luminous bunsen flame over the entire surface. Boiling occurred and from thence onwards every effort was made to keep the liquid boiling steadily since the gradual removal of air from the water and from the glass walls made it increasingly difficult to recommence smooth ebullition once it had ceased. This procedure was continued until the evaporation of the water had lowered the level in the cell by approximately $\frac{1}{2}$ in.; the tap between the pump and the cell was then closed and the cell cooled by surrounding it with an ice and water bath. The cell was sealed off by collapsing the constriction and drawing off. If the cell was inverted cautiously after it had cooled to room temperature a distinct "water hammer" effect was produced.

TECHNIQUE FOR USING THE CELL

The cells are prepared for use by immersing them for at least one hour in a bath of finely divided ice. A sheath of ice several millimetres thick is then formed over the whole length of the thermometer well by rapidly cooling from within. A very pure water-ice interface is next obtained by melting a thin layer of pure ice adjacent to the well. The thermometer well is then filled with ice-cold water and the cell is ready for use.

The sheath of ice can conveniently be formed by inserting into the well a closely-fitting metal rod previously cooled by immersing it, for example, in liquid oxygen or nitrogen. Alternatively the sheath can be formed by injecting a stream of carbon dioxide expanded from a high pressure cylinder or by filling the well with a cooled liquid. To melt a thin layer of ice the well is filled with slightly warmed water and in this way the sheath is freed from the well in a few seconds. That the sheath is detached can be verified by giving the cell a sharp rotatory movement and observing if the ice sheath spins. The warm water in the well is replaced by ice-cold water which quickly takes up the triple-point temperature and provides good thermal contact with the thermometer.

RESULTS

The equilibrium temperature of the triple-point cells may be maintained under favourable conditions for at least five days if the cell is continuously and completely immersed in ice. Fig. 3 gives examples of readings obtained with a platinum resistance thermometer in two triple-point cells. A reproducibility of about $\pm 0{\cdot}00005$° C was given by these cells for five days. The small variations recorded in Fig. 3 might be accounted for by changes in the resistance of the

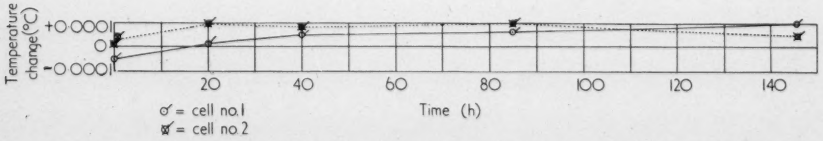


Fig. 3. Platinum resistance thermometer observations in triple point cells. Performance of cells Nos. 1 and 2 over a period of 150 h

thermometer other than those due to temperature, by changes in the measuring bridge or by actual changes in the temperature of the cell. The overall changes recorded in this figure are, however, so small that their source cannot at present be traced with certainty. It must be noted that it was found to be essential to maintain a continuous sheath of ice around the well because if a portion of the sheath melted the temperature no longer remained steady.

Fig. 4 shows the triple-point readings of two platinum thermometers over a period of several months; during this time the thermometers were heated over a range of temperatures from 0 to 100° C for other experiments. A number reconsideration of such factors as the isotopic composition of the water, material of which the cell is constructed and the precise technique for using the cells.

ACKNOWLEDGEMENTS

The work described above has been carried out as part of the joint programme of the National Physical Laboratory and of the Chemistry Research Board. This paper is published by permission of the Director of the National Physical Laboratory and of the Director of the Chemical Research Laboratory.

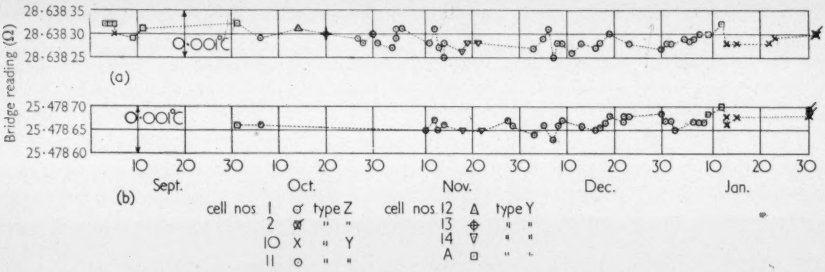


Fig. 4. Triple point of water observations with two platinum thermometers. (*a*) R13, and (*b*) S167

of different cells were employed which were prepared according to the method described here with the exception of cell *A* which had been prepared at the National Bureau of Standards U.S.A. in 1948. Since there are no significant differences between the older cells (No. 1 constructed in 1949 and No. 2 in 1950) and those recently prepared (i.e. Nos. 10, 11, 12, 13 and 14 constructed in 1952), it may be concluded that Pyrex is a suitable material from which to construct the cells and that such cells have a useful life of several years.

The results shown in Fig. 3 on a single preparation of the ice sheath indicate a considerably higher degree of reproducibility than the day to day results given in Fig. 4 where each cell was freshly prepared each day. This suggests that with each fresh preparation of the ice sheath a slightly different equilibrium temperature may be produced. Possibly these slight differences in equilibrium temperature may be due to strain in the freshly prepared ice and, in fact, there is some indication in Fig. 3 that the cells only become really steady 20 h after preparation. However, the difficulties of maintaining a sufficiently high degree of reproducibility in the thermometers and measuring bridge over a long period of time make it uncertain whether or not the fluctuations in readings recorded in Fig. 4 are in fact due to fluctuations in cell temperature. However, when the effects of variations in the measuring instruments (thermometer and bridge) were minimized by comparing a number of cells at one time using a single resistance thermometer it was found that no cell differed from any other by more than 0·0001° C.

As a result of this investigation it is concluded that with reasonable preparations the triple point of water is reproducible to ±0·0001° C with cells prepared as described. To obtain any improvement in accuracy would involve careful

REFERENCES

(1) STIMSON, H. F. *J. Res. Nat. Bur. Stand.*, **42**, p. 209 (1949).
(2) National Physical Laboratory, The International Temperature Scale of 1948 (London: His Majesty's Stationery Office, 1950).
(3) STIMSON, H. F. *J. Washington Acad. Sci.*, **35**, p. 201 (1945).
(4) MOSER, H. *Ann. Phys.* [*Leipzig*], (5), **1**, p. 341 (1929).
(5) WHITE, W. P. *J. Amer. Chem. Soc.*, **56**, p. 20 (1934).
(6) THOMAS, J. L. *J. Res. Nat. Bur. Stand.*, **12**, p. 323 (1934).
(7) BEATTIE, J. A., TZU-CHING HUANG and BENEDICT, M. *Proc. Amer. Acad. Arts Sci.*, **72**, p. 137 (1938).
(8) *Tech. News Bull. Nat. Bur. Stand.*, No. 305, p. 71 (1942).
(9) GLASSTONE, S. *Textbook of Physical Chemistry*, 2nd ed., p. 646 (London: Macmillan and Co. Ltd., 1946).
(10) WICHERS, E., FINN, A. N., and CLABAUGH, W. S. *J. Res. Nat. Bur. Stand.*, **26**, p. 537 (1941).
(11) LA MER, U. K., and BAKER, W. N. *J. Amer. Chem. Soc.*, **56**, p. 2641 (1934).
(12) KIRSHENBAUM, I. *Physical Properties and Analysis of Heavy Water*, 1st ed., pp. 28, 396 and 401 (New York: McGraw-Hill Book Co. Inc., 1951).
(13) EMELÉUS, H. J., JAMES, F. W., KING, A., PEARSON, T. G., PURCELL, R. H., and BRISCOE, H. V. A. *J. Chem. Soc.*, p. 1207 (1934).
(14) PARTINGTON, J. R. *An Advanced Treatise on Physical Chemistry*, Vol. II, 1st ed., p. 281 (London: Longmans, Green and Co., 1951).
(15) LAUG, E. P. *Industr. Engng Chem.* (*Anal.*), **6**, p. 111 (1934).

New books

Experimental nuclear physics (Vol. 1). Editor: E. SEGRÈ. (London: Chapman and Hall Ltd.) Pp. ix + 789. Price 120*s*.

The separate fields of nuclear physics have been presented in a variety of ways, notably in individual monographs on special subjects and in periodical reviews, some for specialists and others for non-specialists, containing authoritative articles of moderate length. Professor Segrè has made a different approach, along the lines of the celebrated Handbuch der Physik.

This first volume comprises five treatises by well-known authorities: Detection methods (H. H. Staub); Passage of radiations through matter (H. A. Bethe and J. Ashkin); Nuclear moments and statistics (N. F. Ramsey); Nuclear two-body problems and elements of nuclear structure (N. F. Ramsey); and Charged particle dynamics and optics, Relative abundances of the elements, Atomic masses (K. T. Bainbridge). The series is intended for serious students and for research workers, to enable them to "catch up" with the progress made in the last few years.

The treatment is advanced and mostly presupposes substantial acquaintance with nuclear physics; it will suit postgraduate students. The quality of the content is excellent, as would be expected from these authors.

It is unfortunate that the main part of Professor Staub's contribution had to be closed at the end of 1950 (a year earlier than the others), since his is one of the most rapidly developing fields. The very thin section on scintillation counters comes as a surprise until this fact is realized. A separate monograph would evidently have saved a year in this case.

The whole series should undoubtedly find extensive use in libraries, and copies will be bought by students who can afford the price and by many of their teachers. However, it is possible that rather few research workers will feel a pressing need to own a book covering each of these five topics at such length. Many might prefer to have one or two sections at lower cost, and to refer occasionally to the library copy for the rest; this is another advantage of monographs.

Nevertheless, all concerned deserve congratulations for a most valuable contribution to the literature of nuclear physics.

T. G. PICKAVANCE

Electron diffraction. By Z. G. PINSKER, translated by J. A. SPINK and E. FEIGL. (London: Butterworths Scientific Publications Ltd.) Pp. xiv + 443. Price 63*s*.

For about ten years after the pioneer experiments of G. P. Thomson there was intense activity in the field of electron diffraction, notably in this country. A relatively inactive period followed and then with the development of the electron microscope as a commercial instrument, came much greater interest in the technique. Today in many countries the interest is greater than ever before and a realization of the limitations of electron diffraction has been accompanied by the exploitation of its advantages when rightly used. It has been possible to read of modern work in most countries, but various difficulties, including those of language, have prevented familiarity with Russian work. Gratitude is therefore due to the translators for providing an English version of Pinsker's book which first appeared in Russian in 1949.

Apart from the detailed description of Russian apparatus, techniques and results, which will be specially welcomed, the book includes a general survey and assessment of what has been achieved by electron diffraction studies. There is possibly a tendency to over-emphasize the importance of Russian contributions, but on the whole the treatment is comprehensive and objective.

This monograph is probably the most complete and up-to-date treatment of the subject at present available. It is well illustrated and appears to have been very well translated, for it reads like an original version. It can be thoroughly recommended to those interested in the principles and application of electron diffraction.

A. G. QUARRELL

Properties of metallic surfaces. (London: The Institute of Metals.) Pp. 368. Price 35*s*.

One of the major problems of science these days is the very great and still increasing tendency towards specialization. Each group of specialists occupies a furrow in a scientific field and one hopes they see something of what is happening in the other furrows. Too often, however, there is the feeling that the furrows have become trenches from which the outlook is very restricted. While we cannot deny the need for more specialization, how can we combat the limited scientific vision that it tends to breed? I think one very fruitful way is illustrated by this recent symposium arranged by The Institute of Metals.

The organizers have chosen for their subject metallic surfaces and a glance at the contents quickly shows that in their wisdom they have cast their net wide. At such a meeting, the specialist must perforce hear discussed work differing very greatly in character from his own, and cross-fertilization becomes possible. That contributions were obtained from physicists, chemists, metallurgists and engineers has ensured that the subject was approached from many different points of view and a wide range of surface effects was covered.

The symposium opens with a survey of specialized optical methods for examining surfaces. Contributions from the field of pure physics include studies of the surface resistance of metals and the crystalline nature of abraded surfaces. On the chemical side, there is an extensive review of the use of radioisotopes for investigating surface reactions, while the important subjects of anodizing and corrosion receive due attention. There is a useful survey of different types of diffusion coatings, but an obvious gap in the chemical papers is the subject of catalysis. The dependence of the mechanical properties of materials on the nature of the surface is well brought out by papers on metal single crystals, glass and the fatigue strength of steel.

Finally, the very practical problem of wear is approached from two viewpoints, that of the engineer and metallurgist who are interested in the effects of machining and lubrication on the wear of moving parts, and that of the physicist and chemist who seek to explain how wear occurs, and what happens to a lubricant on a surface. The papers are followed by a general discussion which, although it is a very mixed bag, provides extremely interesting and informative reading.

The Institute of Metals has clearly met a need for discussion of topics covering a relatively wide field. I trust that this volume will meet with a reception which will encourage the Institute (and other scientific bodies) to arrange many more such symposia in future.

R. W. K. HONEYCOMBE

Automatic digital calculators. By A. D. BOOTH and K. H. V. BOOTH. (London: Butterworths Scientific Publications Ltd.) Pp. vii + 231. Price 32*s*.

The book is intended to introduce new workers to the design, construction and use of automatic digital computing machines. The first three chapters provide a historical summary of developments from 1822 to 1952. The next nine chapters are devoted to their design and construction. The authors cover this field thoroughly and describe devices under development. If, as a consequence, some of the topics do not receive very detailed attention, this is compensated for by the excellent bibliography at the end of the book. The control and the arithmetical units are described and illustrated with block diagrams and circuit elements. The requirements of a storage device are set out and the properties of the magnetic drum, ultrasonic delay and Williams' cathode-ray tube stores are discussed, together with other less common storage devices. A chapter is devoted to input and output equipment.

The last five chapters deal with user aspects. These include the choice of the basic instructions from which are built up the "routines" (sequences of instructions) for carrying out more complex requirements such as interpolation and the calculation of square roots. The assembly of "routines" to form a complete "programme" is briefly mentioned. A final chapter deals with applications. In order to exhibit the scope of digital calculators only one problem of a directly mathematical nature is described, but this is supplemented by such topics as machine "learning" and "intelligence" which the reader should not take too seriously. Unfortunately the bibliography does not provide any alternatives. It seems a pity that the Manchester B-tube is not mentioned as this is one of the few developments in logical design since Babbage.*

R. A. BROOKER
R. L. GRIMSDALE

* KILBURN, T. *Nature* [*London*], **164**, p. 684 (1949).

Flames: Their structure, radiation and temperature. By A. G. GAYDON and H. G. WOLFHARD. (London: Chapman and Hall Ltd.) Pp. xi + 340. Price 55*s*.

The use of fire and flame is one of man's earliest achievements, and yet since prehistoric times it has been surrounded by an aura of mystery. Fire was an essential element in the scheme of alchemy, but only quite recently in its long history has progress been made in understanding its nature. Indeed, the greater part of the work under review is the product of the last decade or so.

This book is an authoritative work written by two active researchers in the field of combustion. The authors confine their attention almost exclusively to stationary flames, as for example, those from the candle and the Bunsen burner, and are not concerned with propagating or explosion flames. The authors' own work forms a substantial proportion of the contents, while relevant work of others is also included and is critically discussed. The emphasis throughout is strongly experimental. The objective of the authors is a general understanding of the physical processes involved without overmuch mathematical theory. As the full title of the book indicates, no attempt is made to give an exhaustive account of all properties of stationary flames. Instead, the authors cover the more physical properties, and mention chemical aspects only when necessary to a better understanding of the physical processes.

The book will be invaluable to those working in the field of combustion and to many who use flames in industrial processes.

E. R. ANDREW

Dielectric aerials. By D. G. KIELY. (London: Methuen and Co. Ltd.) Pp. xii + 132. Price 8*s*. 6*d*.

The author gives a very readable account of the beam forming properties of solid dielectric rods and of thin-walled dielectric tubes. Experimental data derived from measurements are presented, together with a comprehensive review of theoretical attempts to explain the phenomena.

As the author points out, these devices are hardly aerials, but find greater application as *primary radiators*, e.g. as elements of aerial arrays, or as feeds to illuminate lens or mirror objectives. A rod or a tube, being a low-gain, broad-beam device, is unsuited to high resolution microwave radar, or to long-hope microwave repeaters, quite apart from a fundamental sensitiveness to weather conditions and mechanical strains.

The dielectric lens has found useful applications and is possibly the most important *dielectric aerial* as such. It is, however, intentionally omitted. Other forms of sheet radiator are only briefly treated. Dielectric radiators made of expanded or of artificial dielectrics are not reviewed, although the use of these materials leads to first-order changes in the quantitative performance. The important, apparently general, statement on p. 46 that "Beam-widths narrower than about 20° cannot be achieved with a single dielectric rod . . ." then requires modification.

The author points out in detail that the theoretical explanation of the performance of the rod and tube radiators has not reached finality. Rival first-order theories exist, an exact explanation being still lacking. It would seem, however, that if the accounts of the far-field diffraction patterns had been implemented by measurements in the near fields, the amplitude, phase and polarization characteristics then available would have gone far to clarify the theory. These curves or contours would account for the structure of the far field, and would indicate the mechanism by which the "aperture illumination" of the rod or tube is derived from the feed.

A summary monograph of this kind would have benefitted by the inclusion of more extensive references to the existing patent literature.

J. F. RAMSAY

Electron optics. By O. KLEMPERER. (London: Cambridge University Press.) Pp. xii + 471. Price 50*s*.

Glaser's monumental work on electron optical theory (reviewed in the September 1953 issue of this *Journal*) is now followed by the appearance of an equally comprehensive survey of the practical side of the subject. In principle this is the second edition of a monograph published in 1939, but the growth of electron optical applications since then has forced Dr. Klemperer to expand his text more than five times, so that in effect it becomes a completely new work. It is frankly written from the experimental standpoint, and there is not so much fundamental theory as the title might lead one to suppose. A more properly descriptive title would have been "Applied electron optics." The essential theory, the design and the properties of electrostatic and magnetic lenses are described in the first five chapters. Lens aberrations are then treated from a practical point of view, with little theoretical discussion, followed by a detailed exposition of space charge effects. Here, and in the chapter on lenses and emission systems with line focus, the author is describing work in which he has himself played a major part, and his treatment is concise and critical. Elsewhere, as in dealing with simple emission systems, although always clear, he is inclined to be comprehensive in quoting the literature without being sufficiently critical or selective. Indeed, one cannot avoid the

feeling that in such a series of monographs on physics, it is more important to outline the state of theoretical knowledge of principles than to provide a handbook of the present state of practice. From this account the inquiring student will find it difficult to discover what are the urgent problems at present needing solution in electron optics. On the other hand, the book is the most complete account of the art that has yet appeared, and will be the standard reference book on electron optical practice for a long time to come. V. E. COSSLETT

Vacuum tube oscillators. By W. A. EDSON. (London: Chapman and Hall, Ltd.) Pp. xv + 476. Price 60*s*.

This comprehensive book includes chapters on transient response of linear systems, negative resistance oscillators, oscillations in non-linear systems, and feedback systems. This introductory part, of five chapters, thus traces the basic part of the whole treatment and is followed by thirteen other chapters on such matters as resonators, crystal controlled oscillators, practical relaxation oscillators, and on more esoteric subjects including locking and synchronization, frequency multiplication and division, modulation, and automatic frequency control.

The author, who is a Visiting Professor at Stanford University, states in his introduction that the treatment is intended for a graduate (i.e. a post-graduate) course and for working engineers. There seems little doubt that the book would indeed be most useful in this way, particularly for the latter, although the problems given at the ends of chapters often have an undergraduate flavour. There is a list of 352 references, and inspection shows that roughly only one-tenth of these are to pre-1930 work.

It seems unfortunate that the basic, theoretical treatment could not have been separated from the engineering details, thus giving two main parts of which the first might have been useful to undergraduates. Finally, until some enterprising publisher develops a cheap method of production, such as the McGraw-Hill atomic energy series, it is sad to consider that works such as the present excellent book on an important technical subject will in all probability find few private purchasers, at least in Britain. J. D. CRAGGS

Radioactive isotopes. By W. J. WHITEHOUSE and J. P. PUTMAN. (London: Oxford University Press.) Pp. xii + 424. Price 50*s*.

The lack of a suitable manual for the scientific worker who makes practical use of radioactive isotopes has long been felt, and the authors are to be congratulated for filling this great need. The text of the book: (*a*) embraces the general laws concerning nuclear disintegration and the properties of the radiations including the effects of the radiations on matter; (*b*) deals with the production of artificial isotopes in piles and cyclotrons showing the limitations and possible developments of the methods used; (*c*) describes at some length the relative and the absolute measurements of disintegration rate and the principles of the instruments used, the Geiger-Müller counter, owing to its wide application, being described in some detail; (*d*) discusses and exemplifies the application of radiactive isotopes to a variety of problems within a wide range of subjects, this being supplemented by numerous references. A separate chapter is devoted to the manipulation of radioactive materials and to the safety precautions necessary in handling them. The text is completed with tables of the properties of the most important isotopes.

The book is well-composed, clearly written and excellently illustrated. Since it is only intended to be an introduction and not a comprehensive treatise, the limits of the book are purposely imposed by the authors. The list of authorities co-operating in the preparation of the text gives an assurance that the book is free from major errors. S. IGNATOWICZ

Dislocations and plastic flow in crystals. By A. H. COTTRELL. (Oxford: Clarendon Press.) Pp. ix + 223. Price 25*s*.

The International Series of Monographs on Physics generally needs no great commendation from the reviewer. The present volume, however, has some notable virtues which deserve to be stressed. Where a volume has a substantial mathematical background the writer is always in a difficulty in presenting his argument. In the first place the ordinary reader is not primarily concerned with the detailed mathematics and the writer must be selective in his choice of the material to be included; but on the other hand, having decided that some item must be presented, it is necessary to give sufficient detail for the argument to be clear from beginning to end. Too often in books of this kind, the author is content to give the primary assumptions, a little mathematics which is known to the expert and unintelligible to the amateur, and then the answer. Professor Cottrell has avoided this pitfall with remarkable skill and, coupled with his exceedingly clear and concise text, he has produced a volume which makes most interesting reading.

The volume does not deal with every aspect of the subjects chosen for discussion but presents the basic theory of dislocations in relation to slip, yielding, work hardening, annealing and creep of metals. The author has attempted with success to achieve a presentation of interest to the physicist, the metallurgist and the engineer but probably the greatest value of the book will be to the metallurgist. Having read the story told by Professor Cottrell, he need never fear a dislocation again. L. ROTHERHAM

Problems in the advanced theory of functions. Vol. 2. By KONRAD KNOPP. (New York: Dover Publications Inc.) Pp. 138. Price $2.50 (cloth); $1.25 (paper).

This slight book is a collection of problems for use with Knopp's *Theory of functions*. It opens with some additional problems for Part I, but is mainly concerned with the contents of Part II of Knopp's textbook. It therefore contains collections of exercises on singularities, entire and meromorphic functions, periodic functions, analytic continuation, Riemann surfaces and conformal mapping. In the second half of the book solutions of all the exercises are sketched, which should prove a real help to any student reading the subject unaided. The exercises are graded in difficulty and seem to cover the subject well. J. TOPPING

Fundamental processes of electrical contact phenomena. By F. LLEWELLYN JONES, M.A., D.Phil., F.Inst.P. (London: H.M. Stationery Office.) Pp. vi + 66. Price 3*s*.

This report was first prepared in 1948 at the request of the Director of the National Physical Laboratory in order to survey and co-ordinate the experimental information and theoretical analysis available on this subject at that time. A supplement was added in 1949 which is now Part 2, and in the present edition an appendix has been added outlining the work up to 1951. Part 1 deals with the fundamental processes and conceptions, the approach of the members, the separation of the contacts, the mechanism of fine transfer, and sealed contacts. Part 2 deals with measurement of the Thomson coefficient, experimental tests of contact theory by Lander and Germer, and elimination of fine transfer in bismuth metal contacts.

Notes and comments

Bicentenary competition

The Royal Society of Arts will reach its bicentenary in March 1954, and with this in mind its Council is arranging a competition which will focus attention upon the future. The Society, accordingly, offers prizes totalling £500, the largest being £250, for conceptions of life on this planet in the year 2000, and forecasts (in visual or written form) are invited of the future developments which may be looked for in some particular aspect of life related to arts, manufactures and commerce, the field of the Society as defined in its full and original title. For example, a competitor might give ideas of what either transport, housing, food or clothing may be like in 2000. The chief criterion in assessing the entries will be originality.

The general subject of the competition is: "The practical aspects of life on this earth in the year 2000" and individual competitors are required to submit their forecasts of a single aspect of life at that period. Full terms and conditions relating to the competition, together with registration forms, may be obtained from the Secretary, Royal Society of Arts, John Adam Street, London, W.C.2. Registration forms must be completed and returned, together with an entry fee of 1*s*., by 15 February, 1954, and the actual competitive material submitted by 30 June, 1954.

Methods of recording sound

An international meeting on methods of sound recording and their applications in different fields will be held in Paris from 5 to 10 April, 1954. The congress, which is being organized by the Société des Radioélectriciens, will be divided into four sections: Mechanical recording, Photographic recording, Magnetic recording, and Development and standardization problems. An exhibition, comprising the latest types of apparatus together with instruments of historical interest, will also be held.

Further details may be obtained from the Société des Radioélectriciens, 10 Avenue Pierre Larousse, Malakoff, Seine.

Second Radioisotope Conference

It has been announced that the Second Radioisotope Conference will be held in Oxford during the week 19–23 July, 1954. The papers read will deal with new techniques and results obtained using radioisotopes in medicine, biology, agriculture, chemistry, physics, and in technology and industry. To enable adequate time to be devoted to descriptions of new techniques and the experimental uses of radioisotopes in medicine, no papers which give the results of established uses of radioisotopes in medicine will be included. An exhibition of instruments and techniques of interest to radioisotope users will also be arranged.

Details of the arrangements may be obtained from the Atomic Energy Research Establishment, Harwell, Didcot, Berks.

ASTM–IP Petroleum Measurement Tables

The Metric Edition of the ASTM–IP Petroleum Measurement Tables, having 452 pages, has now been published. These tables are now available in British, United States and Metric Editions, and have been constructed upon the same basic data for the coefficients of expansion, etc. In order to meet the language needs of countries using the metric system, the introductory and explanatory matter in the Metric edition has been printed in English, French, and Spanish.

The tables are published by the Institute of Petroleum, 26 Portland Place, London, W.1. Price 55*s*.

Solution of a non-linear differential equation

A report by Mr. H. Goldenberg (Technical Report Z/T88) has recently been published by The Electrical Research Association, Thorncroft Manor, Dorking Road, Leatherhead, Surrey. Price 3*s*. The report provides the general solution of a special case of a non-linear differential equation.

The equation:

$$y(dy/dx) + a = bxy$$

arose in calculations which were made to determine the shape of the current wave in a circuit-breaker before current zero and when there was capacitance in parallel with the arc gap. The solution given can also apply to calculations on a variety of other subjects.

Journal of Scientific Instruments

Contents of the January issue

ORIGINAL CONTRIBUTIONS

Papers

An electromagnetic method of measurement of density or specific gravity of liquids. By K. R. Honick.
The performance of dried and sealed mica capacitors. By G. H. Rayner and L. H. Ford.
A time-marker for electrocardiography. By M. A. Bullen and L. A. Daynes.
A bed level indicator for detecting the boundary of a body of water. By M. J. Wilkie and R. F. J. King.
Pressure chamber for micro-optical observations. By C. Turner.
A small adjustable X-ray crystal monochromator. By J. Fortey and E. Cohen.
A vibrating needle electrometer. By Y. L. Yousef and R. Kamel.
A new form of anomaloscope. By C. R. Forshaw.
Liquid manometers with high sensitivity and small time-lag. By F. A. MacMillan.
Pneumatic gauging applied to the measurement of the bore of tube. By R. Chittleburgh, E. F. Powell and G. F. Morton.
A new method of improving furnace temperature control. By W. James.

LABORATORY AND WORKSHOP NOTES

An electronically operated switching device for the elimination of thermal voltage drifts in a recording Wheatstone bridge. By W. R. Ward and W. F. Maddams.
A specimen screening aperture for the electron microscope. By R. S. Page and A. W. Agar.
A spring-loaded constant stress device. By A. J. Kennedy.
Anti-vibration mount for a sensitive instrument in a boiler house. By O. P. T. Kantorowicz.
A remote valve for dispensing colloidal ^{198}Au solutions. By L. E. Preuss.
A test cell for the measurement of the conductivity of butter and margarine. By J. H. Prentice.
Some improvements in the Hadding tube. By G. B. Mitra and D. M. Chackrabortty.

NOTES AND NEWS

New instruments, materials and tools
Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique).
All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2*s*. 6*d*. including postage.)
REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s*. each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

SPECIAL ARTICLE

Ionization processes in the electrical breakdown of gases*

By Professor F. LLEWELLYN JONES, M.A., D.Phil., F.Inst.P., Department of Physics, University College of Swansea

In elucidating the phenomena of breakdown it is necessary first to find the general mechanism by which the various processes of ionization interact to produce the necessary amplification of current, and then to identify the particular ionization processes themselves. It was the exact prediction, both of the form of the growth of ionization and of the breakdown potentials, when the spark parameter pd was less than about 200 (mm of mercury × cm), that first established the validity of the Townsend theory based on the continuous development of ionization by primary and secondary ionization processes; the main characteristics of this theory are given. The problem of breakdown at higher pressures when pd is $\simeq 760$ (mm of mercury × cm) is next discussed, and recent experimental and theoretical work on the growth of ionization currents and breakdown is described, from which it is concluded that the general Townsend mechanism can also account for breakdown in this range of the spark parameter. This mechanism is then discussed in relation to other properties of breakdown involving the spatial-temporal development of currents, breakdown under impulse voltages, under non-uniform fields, and, finally, in highly compressed gases at very high voltages.

LIST OF SYMBOLS

I_0, initial photoelectric current from the cathode;
I, ionization current in the gas;
d, electrode separation;
d_s, electrode separation at breakdown;
p, gas pressure;
α, primary ionization coefficient due to electrons; $\alpha n dx$ is the number of ionizations produced by n electrons in moving a small distance dx along the electric field;
β, secondary ionization coefficient due to positive ion collisions with gas molecules; $n\beta dx$ is the number of ionizations produced by n positive ions in moving a small distance dx along the electric field;
γ, secondary ionization coefficient due to impact of positive ions on the cathode; γn is the number of secondary electrons emitted from the cathode due to the impact of n positive ions;
δ, secondary coefficient representing the photoelectric effect at the cathode due to radiation produced in the gas by the ionization current; $\delta n dx$ is the number of electrons produced from the cathode when n electrons move a small distance dx in the direction of the electric field;
ϵ, secondary coefficient representing emission from the cathode due to impact of excited atoms; $\epsilon n dx$ is the number of electrons emitted from the cathode when n electrons move a small distance dx through the gas in the direction of the electric field;
η, secondary coefficient representing the photoionization in the gas; $\eta n dx$ is the number of electrons produced by photoionization in the gas when n electrons move a small distance dx through the gas in the direction of the electric field;
ω, the generalized secondary coefficient which includes all the above secondary processes;
V, potential difference across the electrodes;
V_s, breakdown, or sparking, potential;
μ_{760}, absorption coefficient due to ionizing photons at a pressure of 760 mm of mercury.

* This article is based on a lecture delivered at the Conference on ionization phenomena in discharges, held at the Clarendon Laboratory, Oxford, on 20 July, 1953.

THE MECHANISM AT LOW PRESSURES

$[pd_s \lesssim 200$ (mm of mercury × cm)]

The process of electrical breakdown, that is, the rapid transition of a gas from the insulating to the conducting state, involves the production of an extremely large electron and ion current from the electrons initially produced from, say, an external source. To understand the phenomena we require to find the general mechanism by which the various processes of ionization act and inter-act to produce the required amplification.

A crucial method of doing this experimentally is that introduced at the turn of the century by Townsend himself, namely, the measurement of the spatial growth of an electron current in a gas in a static uniform field. It is well known that in all gases at low pressures current growth follows the relation

$$I = I_0 \exp(\alpha d)/\{1 - [\omega/\alpha][\exp(\alpha d) - 1]\} \qquad (1)$$

In practice (ω/α) is small for most gases, so that the log I, d graph is linear for low values of d then curves upwards as d increases until a value of d_s is reached such that the denominator vanishes. This upcurving is of vital significance; the current I approaches infinity asymptotically

when $$1 - [\omega/\alpha][\exp(\alpha d) - 1] = 0 \qquad (2)$$

When the initial ionization is maintained, the ionization current can increase with time. The gas no longer insulates; equation (2) is then the spark breakdown criterion relating the spark parameter $p \times d_s$ and spark voltage V_s giving Paschen's law for those ionization processes depending upon electron and ion energies. The current I following equation (1) is called the *pre-breakdown current*.

It was the exact prediction both of the experimental growth curve and of the breakdown potentials, when the spark parameter (pd) was less than about 150 (mm of mercury × cm), that established the validity of the Townsend theory.

SECONDARY PROCESSES

Consequently, it follows that the breakdown mechanism at these lower pressures is controlled by the primary and secondary ionization coefficients α and ω/α, which for any gas can be measured. To understand the breakdown, therefore, it is necessary to understand the physical significance of the

coefficients α and ω/α in terms of collisional and impact processes involving electrons, atoms, photons, ions and the electrodes.

A number of quite different secondary processes can produce a growth of the ionization current which follows almost exactly the same form as equation (1), and this makes the precise interpretation of the coefficient ω/α difficult in any given case. The possible processes are:

in the gas:
(i) ionization of gas atoms in collision with positive ions (β-effect), and
(ii) photoionization (η-effect), but only under certain conditions;

at the cathode: secondary emission of electrons due to incidence of
(iii) positive ions (γ-effect),
(iv) photons (δ-effect), and
(v) excited atoms (ϵ-effect).

It can be shown, however, that the Townsend theory can be generalized to include all these secondary processes. For,

$$\omega = \beta + \alpha\gamma + \delta + \epsilon + \eta \qquad (3)$$

Although the shape of the ionization growth curve cannot itself decide which of these secondary processes predominate in any given case, other characteristics of the processes do this.

At low pressures the experimental evidence indicates the predominance of processes of cathode emission. Such emission, unlike gas ionization, depends both on the state of the cathode surface as well as on the energies of the photons or ions, and it is on this aspect that considerable work has been done in recent years.[1] This shows how the surface state mainly controls the nature of the secondary ionization ω/α, and experiments at low pressures ($\simeq$ cm of mercury) show that the results for uniform fields have their counterpart in non-uniform fields.[2]

CHARACTERISTICS OF THE GENERALIZED TOWNSEND THEORY

To summarize, it is of interest to tabulate the characteristics of the generalized Townsend mechanism as applied to a static electric field, because many misconceptions of its nature are still held.

(i) The pre-breakdown ionization current follows the well-known growth equation (1), from which the coefficients α and ω/α can be measured experimentally for any gas and electrode of uniform surface properties. It is important to realize that the primary and secondary processes α and ω/α act *throughout* the gap space; it is wrong to conceive of the secondary process, as it were, coming in at large inter-electrode distances only. Secondary coefficients are, in general, low, but the subsequent amplification of secondary electrons by primary ionization is the physical reason for the great importance of even small secondary effects, leading then to rapid growth both in space and time.

(ii) These data enable the breakdown potential V_s to be predicted according to equation (2).

(iii) At low values of $pd \lesssim 100$ (mm of mercury $\times$ cm) there can be a strong dependence of V_s on cathode surface properties, but this dependence diminishes as pd increases, until at high values of pd it becomes less significant experimentally, provided that no different secondary processes are introduced.

(iv) The cross-sectional area of the ionization current in the gap is that determined by the lateral diffusion of electrons and ions. Calculated areas are in general agreement with those observed at both low and high pressures for low self-maintained currents.

(v) For uniform electrode surfaces and low fields, Paschen's law is obeyed, since the ionization collisional processes depend on the ratio electric field/gas pressure, and not on, say, the field only.

(vi) The formative time lag of breakdown under impulse fields should depend on the nature of the predominant secondary process, the value of the overvoltage and the location of the initiatory electrons.

(vii) Finally, breakdown is the result of a continuous development of current by primary and secondary ionization processes: all secondary processes which conform to the growth equation (1) can occur, and those which predominate in any given case depend largely on the geometry and nature of the cathode.

THE PROBLEM OF BREAKDOWN AT HIGHER PRESSURES

$[pd_s \simeq 760$ (mm of mercury $\times$ cm)]

It has long been accepted that these characteristics are all in accordance with observations at low pressures. Now, the relevant ionization processes are continuous functions of the field and pressure, and it would appear that there is no obvious reason for believing that this general theory would not also apply to other gap geometries or to other pressures such as atmospheric.

Up to 1925 this mechanism was considered to account for a spark in the atmosphere, but the experiments of Rogowski and his associates showed that the formative time lag of breakdown for gaps of the order of cm in the atmosphere under high impulse conditions was only about a μsec, and this was considered in some quarters to indicate that the Townsend mechanism no longer applied because positive ions could not cross the gap to the cathode in such a short time. This is no objection, however, because the generalized Townsend mechanism can also involve a photoelectric action at the cathode which leads to formative times of the order of those observed by Rogowski.

There was, however, another crucial class of experiments which did not appear to support the application of the Townsend theory to the higher pressures. These experiments involved the classical test, i.e. the measurement of the amplification of a small initial electron current in a parallel plate gap under static electric fields up to separations at which breakdown occurred at pressures up to atmospheric. This was carried out by Sanders,[3] Posin,[4] and Hochberg and Sandberg,[5] and their results indicated that the log I, d curves were straight up to the distance at which a spark was observed. An important conclusion was then drawn.[6]

The appearance of a spark before any curvature in the log I, d graph was found, was taken to indicate the *sudden introduction* at the sparking distance of an entirely new mechanism of breakdown not dependent on a secondary ionization process as outlined above: i.e. there was no coefficient (ω/α). Thus, breakdown at these pressures ($\simeq$ atmospheric) appeared to present an outstanding problem.

For these reasons, the spark criterion based on the Townsend mechanism of primary and secondary ionization was in some quarters rejected in favour of an entirely different idea of breakdown based on positive ion space charge and photoionization.[7] These space charges were stated to distort the

field and assist photoionization at the head of an electron avalanche, practically in the path of the current, so as to give, it was stated, a narrow filamentary discharge; the process was supposed to occur extremely rapidly (in distance) just when the gap distance became the sparking distance, because for all smaller gap distances only the Townsend α-process was stated to occur to give a linear log I, d graph.

This view was never put into quantitative form to relate the breakdown potential to measurable ionization processes of electron collision and photoionization which was stated to be essential. For no clear reason, it was stated that the new ideas applied when pd_s exceeded 150 (mm of mercury $\times$ cm), and certainly for a 1 cm gap in the atmosphere $pd = 760$ (mm of mercury $\times$ cm). Further, since breakdown was considered to be a gas process only, the cathode was stated to have no influence on the mechanism.

Since the time this view was first proposed, no experimental evidence has been brought forward to indicate whether this field distortion does in fact occur in the way suggested (theoretical considerations of Zeleny[8] suggested the contrary), or whether photoionization is at all a significant process, or that, if it does occur, it acts in the way suggested leading to a narrow channel and sudden breakdown. In particular, no quantitative analysis of the influence of photoionization on the growth of the electron avalanche was made. Finally, no satisfactory source of the requisite high energy photons in a pure gas has been proposed.

Since, however, the basis of this new view was derived from the previous experimental measurements of ionization currents at high pressures it becomes necessary to analyse that work carefully. When this is done it can be seen that sufficient attention had not been paid either to the requisite scope of those measurements themselves or to their accuracy; in particular the field had not been maintained steady enough to allow the current at the requisite large electrode separations to be measured.

EXPERIMENTAL INVESTIGATION OF BREAKDOWN AT HIGH PRESSURES

$[pd_s \simeq 760$ (mm of mercury $\times$ cm)]

For reasons such as these, the growth of currents in the crucial region just prior to the appearance of a spark in air was investigated at Swansea.[9] The results showed that over a restricted distance the log I, d graphs were linear, as previous workers had found; when however, d was increased to approach the crucial region near breakdown, the graphs were found to curve up, contrary to previous work, and were of exactly the same form as those found for low pressure.

To test the spark criterion equation (2) from these growth of current curves, α was found from the straight parts and ω/α from curved portions, and these values were substituted in equation (2) to find d_s and V_s theoretically. After each set of runs a single but extremely careful determination of V_s was made. These values showed full agreement between the calculated and observed breakdown potentials.

The following conclusions were drawn.

(i) Even at high pd_s (spark in air) equation (1) still describes the growth of ionization currents in uniform static fields involving primary and secondary processes, just as in the case of low pressures when $pd_s \lesssim 200$.

(ii) The mechanism of breakdown itself is precisely that which causes the growth of currents in the pre-breakdown regime, from which it is a continuous development.

(iii) No evidence was obtained for the sudden appearance just at the sparking distance of any new ionization process producing a spark.

SECONDARY PROCESSES IN UNIFORM STATIC FIELDS AT HIGH PRESSURES. PHOTOIONIZATION IN AIR

It remains now to discuss the physical significance of the secondary ionization coefficient ω/α. The experiments, using different cathode surfaces, showed that in air the secondary ionization was at least in part cathode emission due to incidence of ions and photons.

Nevertheless, let us consider the question of photoionization. Since air is a mixture of gases of different ionization potentials there is in this gas a possible process of production of high energy photons, because excitation of nitrogen molecules can produce photons capable of ionizing oxygen. On the other hand, in a simple non-active gas no process exists, other than those involving ions with which the rate of secondary ionization would then depend on the square of the current. Further experiments[10] were therefore carried out in a simple gas, nitrogen, and these showed that the all-important upcurving of the log I, d curve occurs in nitrogen just as in air. Furthermore, since the value of the coefficients α and ω/α, and therefore the sparking distance d_s, were the same for a wide range of values of I_0 the experiments also showed that there was no significant ionization which depended on the square of the current, and therefore that photoionization played no significant role in the breakdown of this simple gas, and also that no significant effect attributable to space charges occurred over that range of currents.

It is now interesting to consider air, in which gas a mechanism for the production of ionizing photons does exist. A detailed mathematical investigation of the growth of current in air, assuming completely inactive electrodes, on the basis of photoionization treated as the only secondary process has very recently been made at Swansea.[11] This shows that photoionization can only lead to breakdown under certain conditions which impose severe restrictions on the values of the various photon absorption coefficients. The physical reason for this is simply that it is only by sufficient back-penetration of the gas by photons that the requisite amplification of the photoelectrons to lead to breakdown can occur: the amplification of current was effected mainly by those photoelectrons and not the photoelectrons produced at the *head* of the main avalanche.

The analysis shows that when, as in any practical case, resonance fluorescence and degeneracy occur, then the absorption coefficient of the photons in air which lead to ionization of the molecules must be small ($\mu_{760} < 10\ \text{cm}^{-1}$). It follows from this low value that the photons make considerable penetration of the gas, so much so that a large number of high energy photons will not only reach the cathode but also spread laterally. In, say, a 1 cm gap in air the cathode itself, when composed of an ordinary metal, might still then have the dominating influence on the mechanisms leading to breakdown. Thus if photoionization is assumed to be the predominating secondary process in air, the current path should be comparatively diffuse; and the assumption that photoionization leads to a narrow filamentary path is not in accordance with quantitative calculation. On the other hand, the secondary process involving positive ions at the cathode is more likely than photoionization to give a narrow filamentary channel.

One other conclusion follows from these recent calculations. They show that qualitative considerations of the action of

photons in any particular case of breakdown can be not only inaccurate but also misleading. For gases where it is known that photoionization takes place, the growth of the ionization current should be calculated for the particular gas and electrode configuration under consideration before it can be concluded that photoionization is, or is not, the significant secondary mechanism.

SECONDARY PROCESSES IN NON-UNIFORM STATIC FIELDS

These considerations are relevant to the case of breakdown in non-uniform static fields at atmospheric pressures, when the point or smaller electrode is the anode, and the cathode is at a great distance. These conditions practically eliminate cathode secondary processes leaving only the gas processes of ionization by positive ions or by photon absorption for consideration. However, as stated above, qualitative discussion of the relative importance of these processes is of no great significance: the question can only be decided by reliable measurement of cross-sections for ionization by positive ions (the recent experiments of Horton and Millest[12] give added grounds for re-opening this question) and by accurate measurement of the pre-breakdown ionization currents and detailed quantitative calculation, as has been done for uniform fields: this has still to be done. Further experimental data are therefore required before it can be definitely ascertained which of these two secondary processes predominates in the case of the positive point discharge. These observations are also relevant to the question of breakdown in long gaps, as these are, generally, instances of non-uniform fields.[17]

THE TEMPORAL GROWTH OF IONIZATION CURRENTS

Another important aspect of electrical breakdown in gases is the formative time lag, which is the time elapsing from the appearance of the initiatory electron to the development of any given breakdown current in the gap. This time lag depends on the various ionization processes involved as well as on the overvoltage.

Detailed calculations, taking into account various secondary ionization processes, have recently been made at Swansea[13] for a 1 cm gap at atmospheric pressure. The value of the current which may be taken as indicative that the gap has broken down in any given case was taken as a parameter, and the time of breakdown calculated for a range of values of that current.

The general result shows that, very near the static sparking potential, the time lag, on the generalized Townsend theory is large $\simeq 10^{-4}$ sec because a very large number of avalanches must traverse the gap before the current builds up to the required value. As the overvoltage is increased, however, the efficiency of the ionization processes increases, and fewer avalanches are then necessary to give the same current. The theory predicts that the time decreases rapidly with overvoltage such that, if only 50% of secondary emission (ω/α) were photoelectric, then the time lag is as low as 1 μsec with an overvoltage as small as 2%. The conclusion previously made that such short times with impulse breakdown were inconsistent with this generalized Townsend mechanism is wrong. There is general agreement with these theoretical results and the recent experimental work of Fisher and Bederson.[14]

Another consequence of these calculations[13] concerns the spatial-temporal development of any given value of ion concentration in the gap. On the basis of the development of ionization by primary (α) and secondary (γ and δ) cathode processes, it can be shown that any given value of ion concentration under an impulse field during ionization growth *first appears near the anode*, and then at subsequent times appears at positions nearer and nearer the cathode. Since the detection or recording of light produced by the ionization current depends on a certain minimum ion or electron density being attained, the complete spark discharge can then appear as a glowing avalanche which crosses the gap *from anode to cathode*; the speed of transit being determined by the *spatial-temporal growth of ionization* in the gap. Very short times of apparent transit of the glow can thus be obtained with even low overvoltages $\simeq 5\%$.

BREAKDOWN IN HIGHLY COMPRESSED GASES, FAILURE OF PASCHEN'S LAW

Another group of breakdown phenomena, that at very high voltages and very high gas pressures, is of particular interest. According to Paschen's law, in order to insulate a high voltage V a conveniently short distance d may be employed provided the gas pressure is correspondingly high. In this way high voltage equipment may be made conveniently small. Recent work, however, with comparatively short gaps ($\simeq$ cm) and very high pressures ($\simeq$ 100 atm) show that the breakdown potential was lower than that expected from Paschen's law.[15] It is now necessary to consider the breakdown mechanism in these circumstances.

In the gaps used at 100 atm the electric field is $\simeq 10^5$ V/cm, and recent work at Swansea[16] has shown that electrons can be extracted at a considerable rate with such field intensities from cathodes in air by a field process, thus increasing the ionization in the gaps and so lowering the breakdown potential. The theory of breakdown can thus be generalized to include this field emission; but this would cause departures from Paschen's law as has been previously observed experimentally by Trump and others at the Massachusetts Institute of Technology.[15] The factor of greatest importance when the field is high is the nature of the cathode surface.

CONCLUSION

In concluding this brief survey of breakdown, it would appear then that the generalized theory of current growth dependent on primary and secondary ionization processes can be applied over a wide range of experimental facts, covering various gases, pressures and gap geometries: the factors which mainly determine the predominant secondary processes in any particular case are the nature of the cathode surface and geometry of the system.

REFERENCES

(1) LLEWELLYN JONES, F., and DAVIES, D. E. *Proc. Phys. Soc.* [*London*], *B*, **54**, pp. 397 and 519 (1951).

(2) LLEWELLYN JONES, F., and WILLIAMS, G. C. *Proc. Phys. Soc.* [*London*], *B*, **66**, p. 345 (1953).

(3) SANDERS, F. H. *Phys. Rev.*, **41**, p. 667 (1932).

(4) POSIN, D. Q. *Phys. Rev.*, **50**, p. 650 (1936).

(5) HOCHBERG, B., and SANDBERG, E. *J. Tech. Phys. U.S.S.R.*, **12**, p. 65 (1942).

(6) LOEB, L. B. *Fundamental Processes of Electrical Discharges in Gases*, Chapter 9 and p. 407, Ref. 62 (New York: John Wiley and Sons Inc., 1939).

(7) LOEB, L. B., and MEEK, J. M. *Mechanism of the Electric Spark* (Stanford: The University Press, 1941).

LOEB, L. B. *Phys. Rev.*, **81**, p. 287 (1951).

LOEB, L. B. *Brit. J. Appl. Phys.*, **3**, p. 341 (1952).

RAETHER, H. *Z. Angew. Phys.*, **6**, p. 211 (1953).

(8) ZELENY, J. *J. Appl. Phys.*, **13**, pp. 103 and 444 (1942).

(9) LLEWELLYN JONES, F., and PARKER, A. B. *Nature [London]*, **165**, p. 960 (1950); *Proc. Roy. Soc. A*, **213**, p. 185 (1952).

(10) DUTTON, J., HAYDON, S. C., and LLEWELLYN JONES, F. *Proc. Roy. Soc. A*, **213**, p. 203 (1952).

(11) DUTTON, J., HAYDON, S. C., and LLEWELLYN JONES, F., and DAVIDSON, P. M. *Proc. Roy. Soc. A*, **218**, p. 206 (1953).

(12) HORTON, F., and MILLEST, D. M. *Proc. Roy. Soc. A*, **185**, p. 381 (1946).

(13) DUTTON, J., HAYDON, S. C., LLEWELLYN JONES, F., and DAVIDSON, P. M., *Brit. J. Appl. Phys.*, **4**, p. 170 (1953).

(14) FISHER, L. H., and BEDERSON, B. *Phys. Rev.*, **81**, p. 109 (1951).

(15) TRUMP, J. G., CLOUD, R. W., MANN, J. G., and HANSON, E. P. *Trans. Amer. Inst. Elect. Engrs*, **69**, p. 961 (1950).

(16) LLEWELLYN JONES, F., and DE LA PERRELLE, E. T. *Proc. Roy. Soc. A*, **216**, p. 267 (1953).
LLEWELLYN JONES, F., and MORGAN, C. G. *Phys. Rev.*, **82**, p. 960 (1951); *Proc. Roy. Soc. A*, **218**, p. 88 (1953).

(17) NORINDER, H., and SALKA, O. *Ark. Phys.*, **3**, p. 347 (1951); **5**, p. 493 (1952).

ORIGINAL CONTRIBUTIONS

A standard light source of very low intensity based on the Čerenkov effect

By W. ANDERSON, Ph.D., and E. H. BELCHER, M.A., Ph.D., Physics Department, The Royal Cancer Hospital, London, S.W.3

[*Paper received* 22 *June*, 1953]

Standard light sources of very low intensity for use in studies of luminescent effects have been developed using the visible Čerenkov emission from aqueous solutions of β-ray emitting radioactive isotopes. The measured relative intensities of the emission from various β-ray emitters showed agreement with the classical theory of Frank and Tamm for the Čerenkov effect. The absolute intensities observed were not incompatible with the latter theory. Measurements were also made of the spectral distributions of the emission in the visible range; these distributions showed reasonable agreement with those predicted by the theory of Frank and Tamm. The use of standards based on the Čerenkov effect for the calibration of luminescence detectors using photomultiplier tubes, and for the absolute measurement of low intensities of visible light, is discussed.

In studies of luminescent effects with the photomultiplier tube, it is very desirable that a standard light source of known spectral distribution and known absolute intensity should be available. This permits not only daily checks on the sensitivity of the detecting devices used, but also the interpretation of observed intensities in terms of absolute units. Use has been made of the visible Čerenkov emission of aqueous solutions of radioactive isotopes to develop such a source.

The Čerenkov effect was originally observed[1, 2] in transparent materials irradiated by γ-quanta from external sources, and arises in such cases from an interaction of the secondary electrons produced by absorption and scattering processes with the dielectric field of the medium. The emission occurs as a continuum extending from the ultra-violet to the infra-red, the intensity and spectral distribution being determined solely by the energy of the exciting radiation and the physical characteristics of the medium; it should be noted, however, that the form of the spectrum may be modified by absorption of emitted light within the medium. The effect has recently been shown by Belcher[3,4] to occur also in aqueous solutions containing medium and high energy β-ray emitting radioactive isotopes, arising directly in such systems from the interaction between the emitted β-particles and the dielectric. In those solutions of radioactive species studied which emit β-particles of maximum energy greater than 0·5 MeV, excitation processes within the substrate have been shown to account for less than 2% of the observed total emission intensity.

THEORETICAL DISCUSSION

The classical theory of Frank and Tamm[5] predicts that emission of light will take place in a hollow cone about the track of each secondary electron or β-particle at an angle θ to the direction of motion of the electron defined by

$$\cos\theta = 1/\beta n \qquad (1)$$

where β = velocity of electron relative to that of light;
n = refractive index of medium for emitted light.

The angle θ decreases as the electron is slowed down in its passage through the medium, and the emission vanishes at a lower energy given by

$$\beta = 1/n \qquad (2)$$

For aqueous media, this energy threshold occurs approximately at 0·26 MeV.

Frank and Tamm's theory leads to the following expression for the number N_E of visible photons with wavelengths between λ_1 and λ_2 emitted along the track of a single particle of initial energy E.

$$N_E = 2\pi\alpha\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right)\int_{\beta n>1}\left(1 - \frac{1}{\beta^2 n^2}\right)dl \qquad (3)$$

where $\alpha = 2\pi e^2/hc$, the fine structure constant;
l = total range of particle within the medium.

Belcher[3,4] has evaluated this integral graphically for an aqueous medium from a knowledge of the range-energy function for electrons in water. In the case of a β-ray emitting radioactive isotope in solution, the energy distribution of the emitted β-particles must be taken into account. If p_E is the probability of release of β-particles with energy between the limits of E and $E + dE$, the average total intensity I_β of Čerenkov emission per disintegration is given by

$$I_\beta = \int_0^{E_{max}} p_E \,.\, N_E \,.\, dE \qquad (4)$$

In Table 1, values of I_β calculated from expression (4) for the spectral range 3000–7000 Å are given for three pure β-ray emitting species. The spectral limits were chosen to cover the spectral range of the RCA type 1P21 photomultiplier tube used in the experimental studies.

Table 1. *Average intensity of Čerenkov emission per β-particle*

Radioactive isotope	*Half-life*	*Maximum β-energy*	*Quanta per β-particle between 3000–7000 Å*
^{204}Tl	2·7 years	0·775	5·0
^{32}P	14·3 days	1·69	51
$^{90}Sr + {}^{90}Y$	25 years	2·2 + 0·53	47

In Fig. 1, the spectral distribution of the emission predicted by equation (3) is shown. The form of this distribution should theoretically be independent of the energy of the exciting radiations, provided, of course, that the latter exceed the threshold for the effect in the medium concerned.

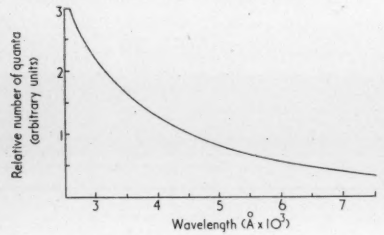


Fig. 1. Theoretical spectral distribution of Čerenkov emission

EXPERIMENTAL

Preparation of light sources. The requirements of a radioactive isotope to be used as a standard light source by virtue of its Čerenkov emission are long half-life, emission of high energy β-particles with no associated γ-radiation, availability at high specific activity, stability in aqueous solution, and freedom from self absorption effects in the spectral range to be studied. These requirements are well satisfied by ^{204}Tl, and also by ^{90}Sr in equilibrium with its daughter isotope ^{90}Y. Studies of the intensity and spectral distribution of the emission from aqueous solutions of both these radioactive isotopes at concentrations suitable for use as standards in studies of luminescent effects are presented below. ^{32}P, although of relatively short half-life, is otherwise suitable, and moreover, is readily available in solutions of accurately known specific activity and has previously been studied in this connexion.[3,4] It has therefore also been included in these studies.

The light sources were prepared as 5 ml, samples of solutions of suitable activity sealed into thin-walled cylindrical glass phials 15 mm in diameter and 0·5 mm in wall-thickness. The radioactive isotopes used were supplied in the following chemical forms:

^{204}Tl as Tl_2SO_4 in dilute H_2SO_4

$^{90}Sr + {}^{90}Y$ as $SrCl_2 + YCl_3$ in 0·1 N HCl

^{32}P as Na_2HPO_4 in normal saline.

The activities of these solutions were measured in substandard liquid Geiger–Müller counters,[6] themselves calibrated by 4π β-counting methods. These activities are considered not to be in error by more than $\pm 5\%$. Solutions of lesser activity were prepared by dilution with distilled water.

Measuring equipment. The measuring system used to study the visible emission from these standard sources is shown diagrammatically in Fig. 2. It comprised a liquid

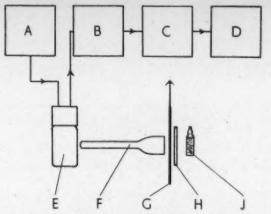


Fig. 2. Diagram of measuring system

A, power unit; *B*, linear amplifier; *C*, pulse discriminator; *D*, counting rate meter; *E*, photomultiplier; *F*, light guide; *G*, shutter; *H*, filter; *J*, source.

nitrogen-cooled nine-stage photomultiplier tube (RCA type 1P21 or Mazda type 27M1), coupled through a cathode follower to a linear amplifier type 1008 operating at its full gain of 16 000 times, and thence to a counting rate meter type 1037A. The input discriminator circuit in the latter unit was set to operate at 10 V and the paralysis time at 10 μsec. The photomultiplier was operated at a potential of 1000 V derived from a power unit type 1082A. The light sources were mounted in a small clip close to one end of a short perspex light guide providing an optical coupling to the photomultiplier. When necessary, filters could be interposed between source and light guide. In the absence of any source the background counting rate of the equipment was of the order of 1 count/sec, this being mainly due to the thermal emission of the cooled photomultiplier.

Measurement of relative intensities of emission. Figs. 3, 4, and 5, are graphs of total counting rate (counts/sec) against specific activity (μc/ml.) for the three species studied. All these measurements were made with a photomultiplier tube type 1P21. The counting rates have been corrected for

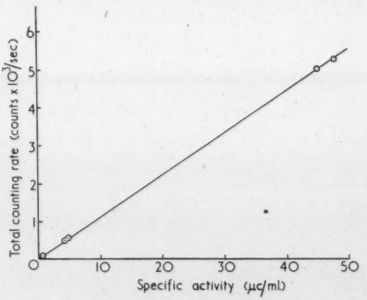


Fig. 3. Variation of total counting rate with specific activity: ^{204}Tl sources

losses due to the finite resolving time of the counting equipment (10 μsec), and also for background effects arising from thermal emission in the photomultiplier and from luminescence induced in the light guide by β-particles escaping from the samples. The latter correction was determined by repeating each measurement with a thin sheet of black paper interposed between the sample and the light guide; in no case did it amount to more than 5% of the observed total counting rate.

The linear relationship between counting rate and specific activity observed in the cases of ^{204}Tl and ^{32}P (Figs. 3 and 5) confirms both the linearity of response of the measuring

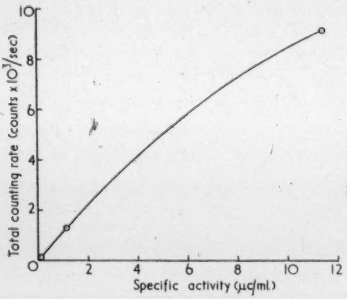


Fig. 4. Variation of total counting rate with specific activity: ^{90}Sr + ^{90}Y sources

equipment in the intensity range studied and the theoretical prediction of direct proportionality between intensity of Čerenkov emission and specific activity for a given β-particle emitter.

The results for (^{90}Sr + ^{90}Y) (Fig. 4) are anomalous, however, in that a fall in efficiency is observed at the highest specific activity studied (11·4 μc/ml.). These results are shown tabulated in Table 2.

Table 2. *Experimental results with ^{90}Sr + ^{90}Y sources*

Sample	*Specific activity* (μc/ml.)	*Corrected total counting rate* (counts/sec)
Sr C	0·114	140 ± 20
Sr B	1·14	1320 ± 200
Sr A	11·4	9200 ± 200

The errors quoted in this and later tables are r.m.s. values based on the maximum ratemeter scale reading errors for both total counting rate and background. These data might suggest a departure from linearity of response in the measuring equipment between 5000 and 10000 counts/sec, but further studies with filters interposed between source and light guide showed them to be attributable rather to self absorption of emitted light in the near ultra-violet region of the spectrum in the case of the more concentrated samples. Self absorption effects of this nature would, of course, increase in magnitude with concentration. That they should occur in the case of the ^{90}Sr + ^{90}Y samples used in these studies is not surprising, since the active solution supplied was derived from aged fission products, and contained appreciable amounts of unidentified inactive material.

In order to investigate the temperature coefficient of emission, sources were warmed gently in a water bath and then allowed to cool to room temperature in the measuring system. No detectable change in counting rate occurred with any of the radioactive isotopes studied during a 20° C change in temperature.

To compare the relative counting rates observed with the various radioactive species studied with the relative intensities predicted by Frank and Tamm's theory and tabulated in Table 1, corrections should strictly be applied for luminescence induced in the glass walls of the phials, and also for luminescence due to excitation of the aqueous substrate. Both of these effects tend to increase the observed intensities. Correction should also be made for the reduction in intensity that occurs because of the escape of β-particles through the walls of the phials. However, with phials of the type and dimensions used in these studies, all the above effects will be small within the β-energy range under consideration, and, moreover, the first two and the third are seen to act in opposite senses. The overall corrections are estimated to be less than 5% of the total counting rates. To a first approximation, therefore, they can be ignored. In Table 3, the total counting rates per μc/ml. for the three species studied, derived from Figs. 3 and 5 and from the initial linear region of Fig. 4, are compared with the theoretical relative intensities per β-particle given in Table 1. The counting rates have been

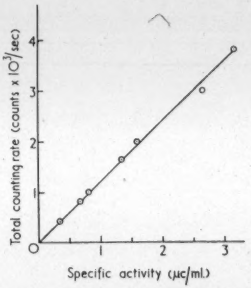


Fig. 5. Variation of total counting rate with specific activity: ^{32}P sources

multiplied by an arbitrary factor so that the value for ^{32}P agrees with theory. Excellent agreement is obtained between theory and experiment in the case of the other two radioactive species.

Table 3. *Relative counting rates with various sources*

Radioactive isotope	*Total counting rate per μc/ml. counts/sec*	*Observed relative intensity*	*Calculated relative intensity*
^{204}Tl	122	4·95	5
^{32}P	1260	**51**	51
^{90}Sr + ^{90}Y	1180	48	47

Measurement of absolute intensities of emission. With regard to the absolute intensities of emission observed, no accurate comparison with Frank and Tamm's theory is possible because of uncertainty regarding the optical efficiency of the experimental arrangement and the quantum efficiency of the photomultiplier cathode, but an approximate comparison is attempted below.

The mean quantum efficiency over the spectral range

3000–7000 Å of the type 1P21 photomultiplier used in these studies was approximately 0·01, whilst the optical efficiency of the experimental arrangement used, including the perspex light guide, was about 0·03. Hence the rate of release of electrons at the photomultiplier cathode due to a 5 ml. sample containing 1 μc/ml. of ^{32}P should according to Frank and Tamm's theory, be

$$5 \times 3{\cdot}7 \times 10^4 \times 51 \times 0{\cdot}01 \times 0{\cdot}03 \text{ electrons/sec}$$

or 2840 electrons/sec.

In order to compare this figure with the observed counting rate, a correction must be applied for those electrons liberated at the photocathode which fail to reach the first dynode and so are not recorded. Morton and Mitchell[7] state that in a type 1P21 photomultiplier operating at 100 V per stage and with a uniformly illuminated photocathode, only 66% of the photoelectrons reach the first dynode. A correction is also necessary for the fact that the counting rate meter is set to record only those pulses which have an amplitude greater than 10 V at its input. Extrapolation of the curve of observed counting rate against pulse discriminator setting indicates that under normal working conditions 80% of the pulses appearing at the photomultiplier output are counted.

These corrections lead to a final figure of $2840 \times 0{\cdot}66 \times 0{\cdot}8$ or 1500 counts/sec for the expected counting rate from a 5 ml. sample containing 1 μc/ml. of ^{32}P. This figure may be compared with the observed value of 1260 counts/sec for such a sample. In view of the uncertainties involved, this result is not incompatible with the classical theory.

Measurement of spectral distribution of emission. The emitted light from the samples used as standard light sources was of too low intensity for direct spectrographic measurement, but the emission from solutions of higher specific

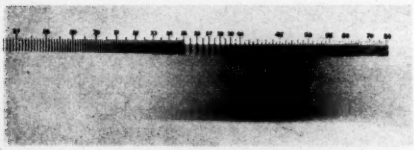


Fig. 6. Spectrum of emission from aqueous solution of $(NH_4)_2HPO_4$ containing ^{32}P

activity can be recorded spectrographically. Fig. 6 shows the spectrum of the emission from a solution of ^{32}P as concentrated aqueous $(NH_4)_2HPO_4$, of initial specific activity 5·25 mc/ml. The active solution was contained in a cubic perspex cell of edge 5 cm placed close to the slit of a small quartz spectrograph by Hilger and Watts, Ltd. The spectrum was obtained after 7 days exposure with a slit width of 1 mm using a Type Oa/F scientific plate by Kodak, Ltd. The image on the original extends from 8000 Å to about 3200 Å, the limit of perspex transmission.

An attempt was also made to investigate the spectral distribution of the emission from the standard light sources, by interposing "Spectrum" series filters by Ilford Ltd., between the sample and the light guide in the photomultiplier measuring system described above. Background measurements were made in every case with a thin sheet of black paper between the sample and the filter, in a similar manner to that described in the previous section. Samples of each of the three radioactive species were investigated in this manner. The samples were of sufficiently high specific activity to give a measurable counting rate with any of ten filters covering the whole visible spectrum. Their specific activities were:

^{204}Tl — 410 μc/ml.
$^{90}Sr + {}^{90}Y$ — 11·4 μc/ml.
^{32}P — 114 μc/ml.

The results of these studies are summarized in Table 4 in which are tabulated the relative counting rates obtained with each of the ten filters for each of the three samples. To obtain these comparative values, the series of observed counting rates for each sample, corrected as described in the previous section for counter resolving time and background effects, have been multiplied by an arbitrary factor to give agreement between the three samples for the counting rates with the deep violet filter. Also included in the table are the calculated relative counting rates for the ten filters based on the theoretical spectral distribution of Fig. 1, and on the published transmission curves of the filters and spectral sensitivity curve of the photomultiplier tube.

No significant difference is observable within the visible spectrum between the spectral distributions of the three samples studied; this result supports the prediction that the distribution is independent of the energy of the exciting radiation. Furthermore, in view of the uncertainties in the filter transmission and photomultiplier spectral sensitivity data, the theoretical and experimental distributions show a reasonable agreement.

DISCUSSION

The foregoing studies have clearly demonstrated that aqueous solutions of β-ray emitting radioactive isotopes can be used as low intensity light sources by virtue of their visible Čerenkov emission. The emission has been shown to be linearly proportional to specific activity for a given radioactive species, provided that self absorption effects are not

Table 4. *Spectral distribution of emission from various sources*

	Filter	*Relative corrected counting rate* ^{204}Tl	$^{90}Sr + {}^{90}Y$	^{32}P	*Calculated relative intensity*
600	deep violet	**34 ± 1·2**	**34 ± 6·7**	**34 ± 5·5**	**34**
601	violet	102 ± 1·2	114 ± 6·7	109 ± 5·5	175
602	blue	48 ± 1·2	54 ± 6·7	49 ± 5·5	39
603	blue-green	14 ± 1·2	16 ± 0·80	15 ± 0·55	28
604	green	11·5 ± 1·2	12·8 ± 0·80	11·6 ± 0·55	15
605	yellow-green	3·8 ± 0·12	5·1 ± 0·80	3·4 ± 0·55	5·4
606	yellow	2·2 ± 0·12	3·2 ± 0·80	1·75 ± 0·55	3·6
607	orange	3·2 ± 0·12	4·6 ± 0·80	1·26 ± 0·080	4·4
608	red	0·51 ± 0·12	1·5 ± 0·80	0·43 ± 0·080	0·67
609	deep red	0·083 ± 0·02	0·86 ± 0·80	0·21 ± 0·080	0·15

significant, and to be dependent only on the physical nature of the medium. Temperature effects are negligible, and a solution of a β-ray emitter of long half-life such as ($^{90}Sr + {}^{90}Y$) behaves effectively as a light-source of constant intensity. Such standard sources have been found invaluable for daily checks on the sensitivity of photomultiplier equipment used in studies of luminescent effects. As an illustration of their use, it was discovered with the aid of a $^{90}Sr + {}^{90}Y$ standard that variations in sensitivity of the order of 20% were occurring in the liquid nitrogen-cooled photomultiplier equipment used in the experimental studies already described. Constant reference to this standard showed that the variations arose from three causes: multiplier fatigue, condensation of moisture on the optical system, and drift in the electronic equipment. As a result, they were largely eliminated.

Again, use has been made of the linear relationship between emission intensity and specific activity to check the linear response of the equipment over a wide range of intensities. Fig. 7 shows a curve of calibration for an arrangement using

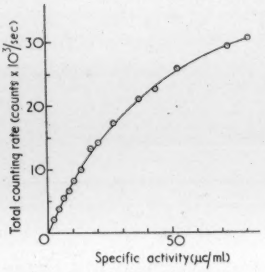


Fig. 7. Calibration curve for measuring equipment based on serial measurements with ^{32}P sources

a photomultiplier tube type 27M1. The curve was obtained by the use of five ^{32}P standards each measured on four different occasions. Their specific activities at the various times of measurement were calculated from a knowledge of the half-life of ^{32}P. Measurements with filters demonstrated that self absorption effects were negligible in the case of these standards, and that the linear relationship between emission intensity and specific activity was maintained over the range of concentrations studied.

It has also been demonstrated that the theory of Frank and Tamm can be used to obtain a theoretical estimate of the absolute intensity and spectral distribution of the visible Čerenkov emission from a standard source of known specific activity. Some caution is necessary, however, in the interpretation of experimental measurements made with such standards in terms of absolute units. In most luminescent processes, the emission of visible light takes place in the form of single quanta randomly distributed in time. In the Čerenkov effect, by contrast, each particle emits a number of quanta along its track, and the emission thus consists of light pulses randomly distributed in time, each pulse having a duration of the order of 10^{-9} sec, and containing several, and in some cases hundreds or thousands of quanta.

Where the detecting device is one that measures integrated light intensity, for example in terms of the direct current drawn by a photomultiplier tube, the discontinuous nature of the Čerenkov emission is no disadvantage. But where the detecting device is of a type that measures counting rate, as is the case in the equipment used for the studies described in this report, then observed counting rates can only be used as measures of emission intensity provided that the pulse amplitude distributions at the photomultiplier output are in all cases comparable. The pulse amplitude distribution appropriate to continuous light is that distribution corresponding to the emission of single electrons at the photomultiplier cathode. That a similar distribution shall occur in the case of the discontinuous Čerenkov emission requires that pulses containing two or more electrons shall not be produced in significant numbers at the photomultiplier cathode. This in turn requires that either the β-particle energy, or the optical efficiency of the system, or the quantum efficiency of the photomultiplier cathode shall be sufficiently small. These conditions were in fact satisfied in the experimental studies described above, but would not necessarily be so satisfied in the case of a system of high optical efficiency using a sensitive photomultiplier tube.

With these reservations, standard light sources of the type described provide a simple method for estimating the absolute intensity of emission in studies of low intensity luminescence. Whilst it has not yet been possible accurately to confirm the correctness of Frank and Tamm's theoretical expression for absolute Čerenkov emission intensity, this expression is certainly correct within an order of magnitude, and the agreement with theory of the relative intensities observed with radioactive species of different β-particle energies does not indicate any wide deviation from the classical theory. It is hoped in the near future to investigate the absolute intensity of Čerenkov emission with greater accuracy.

Note added in proof. Greenfield, M. A., Norman, A., Dowdy, A. H., and Kratz, P. M. (*J. Opt. Soc. Amer.*, **43**, p. 42, 1953) have independently studied the spectral distribution of the Čerenkov emission from water irradiated by ^{32}P β-rays and radium γ-rays, using photographic methods. These workers find that in the limited wavelength range studied by them (313–473 mμ) the distributions agree with that predicted by Frank and Tamm's theory, and they suggest the use of Čerenkov emission for spectral sensitivity measurements of photographic emulsions. Their findings are in complete agreement with our own results.

ACKNOWLEDGEMENTS

It is a pleasure for the authors to record their indebtedness to Professor W. V. Mayneord, Director of the Physics Department, for his interest and encouragement throughout these studies. Acknowledgements are also due to the Governors of The Royal Cancer Hospital and to the Medical Research Council who supported the project. Valuable assistance in the experimental work was given by Mr. J. O. Crookall.

REFERENCES

(1) Čerenkov, P. A. *C.R. Acad. Sci. U.R.S.S.*, **2**, p. 451 (1934).
(2) Čerenkov, P. A. *Phys. Rev.*, **52**, p. 378 (1937).
(3) Belcher, E. H. *Proc. Roy. Soc. A*, **216**, p. 90 (1953).
(4) Belcher, E. H. *Nature* [*London*], **170**, p. 571 (1952).
(5) Frank, I., and Tamm, I. *C.R. Acad. Sci. U.R.S.S.*, **14**, p. 109 (1937).
(6) Veall, N. *Brit. J. Radiol.*, **21**, p. 347 (1949).
(7) Morton, G. A., and Mitchell, J. A. *R.C.A. Rev.*, **9**, p. 632 (1948).

Accuracy limitations of β-ray thickness measurement

By L. MANDEL, B.Sc., Ph.D., A.Inst.P., Central Instrument Section, Imperial Chemical Industries Ltd., Welwyn

[*Paper received* 23 *July*, 1953]

The output fluctuations of β-ray thickness gauges and the limitations they impose on the accuracy of measurement are investigated. The results, based partly on theoretical work and partly on some experiments with ^{204}Tl and ^{90}Sr sources, reveal a simple relation between the shape of the β-ray absorption curve and the region of greatest accuracy. It is shown phenomenologically that the statistical inaccuracies are not improved by the introduction of collimating grids.

Thickness measurement of thin materials based on the absorption of β-rays is becoming an increasingly popular technique and a quantitative description of its distinctive limitations is, therefore, of some interest. It is well known that the output signal of β-ray gauges is usually subject to appreciable random fluctuations, due to the stochastic nature of the radioactive disintegration process. The fluctuations naturally impose a limit on the accuracy of measurement and, in most applications of the instruments in which a reasonably rapid response is called for, it is found that the stochastic inaccuracies are large compared with those arising from other causes. It seems not unreasonable, therefore, to define the achievable accuracy in terms of the output fluctuations. An attempt to relate the two in practice was made by Smith,[1] while Berthold[2] estimated the accuracy of a back-scattering β-ray gauge on this basis.

GENERAL CONSIDERATIONS OF THE ACCURACY OF MEASUREMENT

In the usual method of β-ray thickness measurement the absorbing material of thickness x to be measured is placed between the radioactive source and the β-ray detector (Fig. 1). The mean output signal V from the detector (i.e. the average taken over a long time) is then some function of x, which is characteristic of the source, the detector and the absorbing material, although it does not depend very strongly on the latter if x is expressed in g/cm^2. In general, therefore, we may write:

$$V = V(x) \tag{1}$$

It then follows that small changes in the thickness Δx will be related to small changes in the mean output signal ΔV by:

$$\Delta V \simeq (d/dx)V(x)\Delta x \tag{2}$$

Now the instantaneous output [as distinct from the mean $V(x)$] fluctuates with a certain standard deviation $\sigma(x)$, which is generally small. It seems reasonable, therefore, to take the smallest change in thickness Δx_{min} which is recognizable in single readings of the instrument to be that given by equation (2) when $|\Delta V| = \sigma(x)$. Thus:

$$\Delta x_{min} = -\sigma(x)\Big/\frac{d}{dx}V(x) \tag{3}$$

(The minus sign is inserted to make Δx_{min} positive, since $(d/dx)V(x)$ is always negative.) When measurements are carried out, not by spot readings, but by averaging the output over a certain time or by averaging a series of readings, equation (3) is still valid, provided the standard deviation $\sigma(x)$ is that appropriate to these conditions.

The ratio $x/\Delta x_{min}$ will be denoted by $\beta(x)$. It bears a close analogy to the commonly used signal-to-noise ratio and is a convenient measure of the accuracy achievable with the instrument for any thickness x, under the given conditions. From equation (3),

$$\beta(x) = -\left[x\frac{d}{dx}V(x)\right]\Big/\sigma(x) \tag{4}$$

Now both the functions $V(x)$ and $\sigma(x)$ are largely determined by the nature of the source and of the β-ray detector. For thickness measurement the most commonly used detector is an integrating ionization chamber, whose statistical properties we must therefore examine before equation (4) can be further developed.

THE STATISTICS OF THE INTEGRATING IONIZATION CHAMBER

In this instrument the charged particles are detected by their ionizing action in passing through a sealed chamber. The resulting ions are then collected by the chamber electrodes to which a potential is applied externally (Fig. 1). The

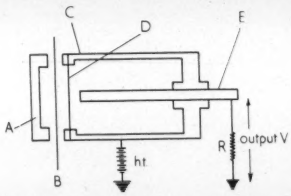


Fig. 1. The arrangement of the ionization chamber and source in β-ray thickness measurement

A, source holder; *B*, absorbing material; *C*, ionization chamber; *D*, chamber window; *E*, cathode.

chamber with its load resistance R forms a charge integrating RC network.

The rate of ion formation—and hence the output potential V developed across R—will depend on the energy spectrum of the incident electrons. Suppose that electrons with energies between E and $E + dE$ enter the chamber at a mean rate $\nu(E)dE$ per sec, and that each gives rise to $\mu(E)$ ions on the average. The actual numbers will, of course, fluctuate about their respective means. Let ϵ be the charge carried by each ion.

Then the average rate of ion collection $\bar{n}$ is given by:

$$\bar{n} = \int_0^\infty \mu(E)\nu(E)dE \tag{5}$$

provided the applied potential is high enough to prevent

appreciable ionic recombination.[3] The mean potential V developed across the resistance R is then

$$V = R\epsilon\bar{n} \tag{6}$$

Since $\nu(E)dE$ is a measure of the frequency with which $\mu(E)$ ions are formed, it is convenient to introduce the notation:

$$\left.\begin{aligned}\bar{\mu} &= \int_0^\infty \mu(E)\nu(E)dE \Big/ \int_0^\infty \nu(E)dE \\ \text{and} \qquad \overline{\mu^2} &= \int_0^\infty \mu^2(E)\nu(E)dE \Big/ \int_0^\infty \nu(E)dE\end{aligned}\right\} \tag{7}$$

Also let

$$N = \int_0^\infty \nu(E)dE \tag{8}$$

the mean rate at which β-rays enter the ionization chamber. With this notation equations (5) and (6) become:

$$\bar{n} = \bar{\mu}N \tag{9}$$

$$V = R\epsilon\bar{\mu}N \tag{10}$$

Thus, the output voltage is proportional to the mean β-ray collection rate. In general, both $\bar{\mu}$ and N will be functions of the absorber thickness x.

In calculating the standard deviation σ of the output we must take the fluctuations of both $\mu(E)$ and $\nu(E)$ into account. We shall first confine our attention to a narrow energy band between E and $E + dE$. Within this band, the mean number of electrons entering the chamber per second is $\nu(E)dE$, with a distribution variance which we shall denote by v_ν. The mean number of ions formed by each electron is $\mu(E)$ with a distribution variance v_μ. The average rate of ion formation is then $\mu(E)\nu(E)dE$ and the variance will be denoted by $v_{\nu\mu}$. The statistical problem in which one set of particles is the progeny of another set is well known in connexion with the theory of proportional counter action. It is found that the following relation holds between the variances[4]:

$$v_{\nu\mu} = v_\nu\mu^2(E) + v_\mu\nu(E)dE \tag{11}$$

Now the number of β-rays incident per second is distributed in a Poissonian form to a good order of approximation. We may, therefore, write:

$$v_\nu = \nu dE$$

and equation (11) then becomes:

$$v_{\nu\mu} = \mu^2(E)\nu(E)dE\left[1 + v_\mu/\mu^2(E)\right] \tag{12}$$

Let us now examine the form of v_μ. The fluctuations of the number of ions formed are due to two main causes, namely:

(*a*) ionization straggling, arising from the stochastic nature of the ionization process (energy loss by radiation is always small at the energies involved in β-ray thickness measurement);

(*b*) the distribution in path length of the β-rays traversing the chamber.

Consider first the fluctuations under (*a*). For electrons which are brought to rest in the chamber, Fano[5] has shown that $v_\mu = (\frac{1}{4}$ or $\frac{1}{3})\mu(E)$. For higher energy particles which traverse the chamber the ratio $v_\mu/\mu(E)$ is somewhat increased. According to the theories of Landau[6] and Symon[7] (the latter is partly reproduced by Rossi[8]), the spread of the distribution is rather wider than Poissonian by a factor not greater than about 6, in the range of energies generally of interest in β-ray thickness measurement. Hence, due to ionization straggling, $v_\mu < \gamma\mu(E)$, where γ is of the order 6.

Consider now the fluctuations arising under (*b*). The path length distribution will, in general, be a function of the geometry of the arrangement. For a typical set up as shown in Fig. 1, with the assumption of isotropic β-ray emission, very rough calculations show that the fraction (variance of path length)/(mean path length)^2 is certainly less than 10%. Now, for particles which traverse the chamber, the path length is very approximately proportional to the number of ions formed in the chamber. Hence, due to the path length distribution, we may say that $v_\mu < \eta\mu^2(E)$, where $\eta \ll 1$.

On combining the two effects we obtain:

$$v_\mu < \gamma\mu(E) + \eta\mu^2(E) \tag{13}$$

It is clear, therefore, that the quantity in square brackets in equation (12), namely:

$$1 < \left[1 + v_\mu/\mu^2(E)\right] < 1 + \gamma/\mu(E) + \eta$$

is approximately unity if $\mu(E) \gg \gamma$. Since the energy loss is of the order 30 eV for each ion formed,[9] this condition will be satisfied provided we are dealing with β-ray energies above about 1 keV. Now $v_{\nu\mu}$ must be integrated over all energies to give the total variance v_n of the number of ions collected per second. It will, therefore, be permissible to neglect the factor in square brackets in equation (12) to a fair order of approximation. Thus:

$$v_n = \int_0^\infty \mu^2(E)\nu(E)dE$$

and from equations (7) and (8),

$$= \overline{\mu^2}N \tag{14}$$

We have now obtained the mean rate of ion collection, equation (9), and the variance of this rate, equation (14). To derive the standard deviation σ of the output potential developed across the resistance R we must take account of the charge integrating action of the ionization chamber. The effect of a resistance-capacity circuit on the output fluctuations is well known in connexion with the behaviour of rate-meters[10,11] and it is found that the standard deviations of the output signal and of the collection rate are related by the factor $R\epsilon/\sqrt{(2\tau)}$. Thus:

$$\sigma = (\overline{\mu^2}N)^{\frac{1}{2}}\left[R\epsilon/\sqrt{(2\tau)}\right] \tag{15}$$

where $\tau = RC$, the effective time constant of the ionization chamber network (see Fig. 1). The standard deviation σ_T of the output averaged over a time T is simply related to σ by[12]

$$\frac{\sigma_T}{\sigma} = \sqrt{\left[\frac{2\tau}{T}\left(1 - \frac{\tau}{T} + \frac{\tau}{T}e^{-T/\tau}\right)\right]} \tag{16}$$

It must be understood that $\bar{\mu}$ and $\overline{\mu^2}$ are functions of the energy spectrum of the incident β-rays and this spectrum will, in general, depend on the amount of absorbing material x placed between source and chamber. It follows that $\bar{\mu}$, $\overline{\mu^2}$, N, V and σ are all implicitly functions of x.

THE DERIVATION OF $\beta(x)$ FOR THE IONIZATION CHAMBER

We can now determine $\beta(x)$ in equation (4) by using equations (10) and (15) [and equation (16) if measurements

are carried out by averaging the output over a time T]. We then find:

$$\beta(x) = -\sqrt{\left(\frac{2\tau}{R\epsilon}\frac{\bar{\mu}(x)}{\overline{\mu^2}(x)}\right)}\frac{x(d/dx)V(x)}{\sqrt{V(x)}} \tag{17}$$

For making numerical estimates it is convenient to express $\beta(x)$ in terms of the normalized functions $V(x)/V(0)$, $\bar{\mu}(x)/\bar{\mu}(0)$ and $\overline{\mu^2}(x)/\overline{\mu^2}(0)$. If we denote the β-ray collection rate without absorbing material by N_0, equation (17) takes the form:

$$\beta(x) = -\sqrt{(2N_0\tau)}\frac{\bar{\mu}(0)}{\sqrt{\overline{\mu^2}(0)}}\sqrt{\left[\frac{\bar{\mu}(x)}{\overline{\mu^2}(x)}\Big/\frac{\bar{\mu}(0)}{\overline{\mu^2}(0)}\right]}\frac{x(d/dx)V(x)/V(0)}{\sqrt{[V(x)/V(0)]}} \tag{18}$$

In principle, the three unknown functions of x occurring in equation (18) can be found when the energy spectrum of the incident electrons is known. In practice, it is often simpler to derive them from observations of the output voltage $V(x)$ and its fluctuations $\sigma(x)$. Thus, from equations (10) and (15):

$$\sigma^2(x)/V(x) = (R\epsilon/2\tau)\overline{\mu^2}(x)/\bar{\mu}(x)$$

so that

$$\frac{\overline{\mu^2}(x)}{\bar{\mu}(x)}\Big/\frac{\overline{\mu^2}(0)}{\bar{\mu}(0)} = \frac{\sigma^2(x)}{V(x)}\Big/\frac{\sigma^2(0)}{V(0)} \tag{19}$$

It follows that all the functions of x involved in equation (18) are derivable from measurements of the output voltage and its fluctuations. Some experiments to this end were undertaken and will now be briefly described.

EXPERIMENTAL DETERMINATION OF $V(x)$ AND $\sigma(x)$

The geometrical arrangement for the experiments was as shown in Fig. 1 and the associated circuits are given in block diagram form in Fig. 2. The output voltage from the chamber was applied to a vibrating reed electrometer which generated

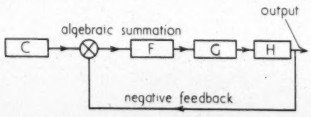


Fig. 2. Block diagram of the measuring circuits

C, ionization chamber; *F*, vibrating reed electrometer; *G*, amplifier; *H*, rectifier.

a proportional alternating potential. This was then amplified, rectified and fed back to the opposite end of the leakage resistance R. In equilibrium, the feedback voltage was equal to the potential $V(x)$ developed across R and could be measured on a potentiometer for different absorber thicknesses x. Measurements were carried out with two commonly used β-ray sources, namely ^{204}Tl, whose β-rays have a maximum energy of $\simeq 0{\cdot}78$ MeV, and ^{90}Sr, containing its daughter isotope ^{90}Y, which emits electrons with energies up to $\simeq 2{\cdot}2$ MeV. The absorbing material was copper, but the ion chamber window and the protective source covering consisted of aluminium foil of mass equivalent slightly below 10 mg/cm^2. The forms of $V(x)/V(0)$ found experimentally are shown in Fig. 3. It was thought more realistic not to correct the curves for the 10 mg/cm^2 additional absorber, since an amount of this order will always be present in practice.

For the determination of $\sigma(x)$, the negative feedback link was broken and, instead, the chamber voltage was approximately balanced against a dry battery connected in series with R. The rectifier output was then applied to an electronic recorder so that the signal fluctuations were directly recorded on a chart. By suitable adjustment of the amplifier gain, it

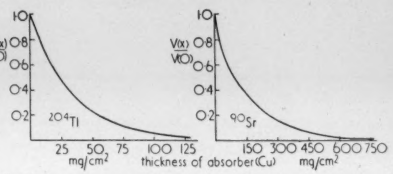


Fig. 3. The absorption in copper of the β-rays from two sources

(Additional absorber $\simeq$ 10 mg/cm^2.)

was possible to maintain the amplitude of the recorded fluctuations approximately constant as the absorber thickness x was changed. $\sigma(x)$ was then inversely proportional to the amplifier gain and could be measured.

Some experimental values of the ratio $\frac{\sigma^2(x)}{V(x)}\Big/\frac{\sigma^2(0)}{V(0)}$ are shown in Fig. 4 for ^{204}Tl and ^{90}Sr β-ray sources. In view of the random character of the signal excursions, accurate

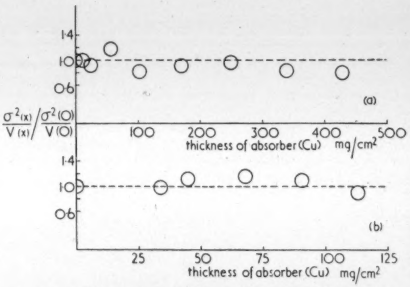


Fig. 4. The variance-to-signal ratio as a function of bsorber thickness (additional absorber $\simeq$ 10 mg/cm^2)

(*a*) Experimental points with a ^{90}Sr source.
(*b*) Experimental points with a ^{204}Tl source.

matching of the fluctuation amplitudes was difficult and the experimental errors are quite large. However, within the limits of these errors, there is no evidence that this ratio varies with x to any appreciable extent over the ranges in question. This conclusion will not necessarily hold for every β-ray source, but it may be taken as a guide.

THE FORM OF $\beta(x)$

Under the stated conditions our experimental results may now be used to simplify equation (18) as follows:

$$\beta(x) = -\sqrt{(2N_0\tau)}\frac{\bar{\mu}(0)}{\sqrt{\overline{\mu^2}(0)}}\frac{x(d/dx)V(x)/V(0)}{\sqrt{[V(x)/V(0)]}} \tag{20}$$

If we denote the ratio $V(x)/\sigma(x)$ (i.e. the signal-to-noise ratio) by $\alpha(x)$, then we have from equations (10) and (15)

$$\alpha(x) = \sqrt{[2N(x)\tau]}\bar{\mu}(x)/\sqrt{\overline{\mu^2}(x)} \tag{21}$$

With this notation (20) now becomes:

$$\beta(x) = -\alpha(0)\frac{x(d/dx)V(x)/V(0)}{\sqrt{[V(x)/V(0)]}} \qquad (22)$$

Thus the accuracy $\beta(x)$ may be determined for any x from the electron absorption curve and the signal-to-noise ratio without absorber.

It is helpful to examine first the general form of $\beta(x)$ to be expected from equation (22). Since $V(x)/V(0)$ and its derivative are finite and non-zero for $x = 0$, $\beta(x)$ is zero at $x = 0$ and increases with increasing x. Further, since $V(x)/V(0)$ and its derivative decrease and tend to zero with increasing x, $\beta(x)$ passes through a maximum and then falls again. It is in the vicinity of this maximum that the β-ray gauge is most suited to thickness measurement. It is worth noting that, when the absorption curve has an exponential form and is given by $V(x) = V(0)\exp(-x/a)$, we find from equation (22) that the greatest accuracy is achieved when $x = 2a$.

The function $$-\frac{x(d/dx)V(x)/V(0)}{\sqrt{[V(x)/V(0)]}}$$

has been computed from the experimental absorption curves of Fig. 3 and the results are shown in Fig. 5. It will be seen that there is a relatively broad region round each maximum where the accuracy is high.

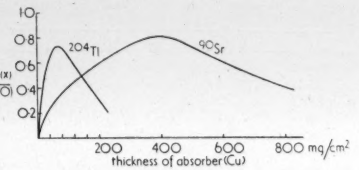


Fig. 5. The dependence of the accuracy on absorber thickness

(Additional absorber ≃ 10 mg/cm².)

For numerical estimates of $\beta(x)$ the factor $\alpha(0)$ must be determined and this involves a knowledge of the ratio $\bar{\mu}(0)/\sqrt{\overline{\mu^2}}(0)$. From the β-ray spectra of ^{204}Tl and ^{90}Sr sources and the known ionization-energy relationship for electrons, $\bar{\mu}(0)/\sqrt{\overline{\mu^2}}(0)$ can be calculated. Rough estimates of the integrals in equation (7) show that this ratio is close to unity for most ordinary sized ionization chambers, so that $\alpha(0)$ is nearly equal to $\sqrt{(2N_0\tau)}$. τ is, of course, the *effective* chamber time constant, when due allowance is made for stray capacities, negative feed-back, etc.

There has been a tendency in the literature on β-ray gauges to ignore the significance of the factor $\sqrt{(2N_0\tau)}$ and to judge the measuring accuracy solely by the shape of the absorption curve. Thus, the introduction of filters[2] and collimators[13] has sometimes been advocated without reference to the resulting decrease in β-ray intensity N_0. Where the measuring accuracy is limited by fluctuations, this is clearly unjustified. We shall examine this question in more detail later.

Finally, it is worth noting that all the foregoing results apply also to a non-integrating detector, such as a Geiger counter, which is used in conjunction with a rate-meter of time constant τ. It is merely necessary to make the substitution $\bar{\mu}(x) = \sqrt{\overline{\mu^2}}(x) = 1$ for all values of x, in the above equations.

THE MINIMUM THICKNESS

From Fig. 5 we note that $\beta(x)$ falls rapidly for thin absorbers and, when the required accuracy and the factor $\sqrt{(2N_0\tau)}$ are specified, the thickness of material which can be measured to this accuracy has a lower limit x_{min}.

It is possible to obtain a very rough estimate of x_{min} from equation (22) by substituting $x \simeq 0$ in all the functions of x. We then find:

$$x_{min} \simeq -\frac{\beta}{\sqrt{(2N_0\tau)}[(d/dx)V(x)/V(0)]_{x=0}} \qquad (23)$$

In particular, for ^{204}Tl,

$$x_{min} \simeq 35\beta/\sqrt{(2N_0\tau)}\ \text{mg/cm}^2$$

and for ^{90}Sr,

$$x_{min} \simeq 85\beta/\sqrt{(2N_0\tau)}\ \text{mg/cm}^2$$

Since β-ray instruments are often in demand for measuring very thin materials, the limitations imposed by expression (23) are certainly undesirable and suggestions have been made for means of lowering this limit of useful measurement. One of these methods has been investigated experimentally and this work will now be briefly described.

EXPERIMENTAL INVESTIGATION OF THE EFFECT OF β-RAY COLLIMATION

In 1951 Syke[13] described a method of collimating the β-rays by the introduction of grids on either side of the material to be measured. Electrons which are appreciably scattered but not stopped by the absorber are then stopped by the grid. The absorbing effect of thin materials is, therefore, exaggerated and the absorption curve becomes steeper for small thicknesses. This increase in slope is, however, obtained at the price of a reduction in the overall β-ray intensity.

In order to investigate the effect of collimation the absorption of β-rays from ^{204}Tl was examined separately with each of three different pairs of grids. The experimental arrangements are shown in Fig. 6 in which the grids have been

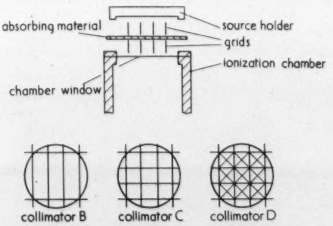


Fig. 6. The geometry of the collimators

labelled *B*, *C* and *D*. Grid *B* had a very coarse mesh and introduced only slight collimation, while grid *D* had the greatest collimating effect.

EXPERIMENTAL RESULTS

The absorption curves obtained with this arrangement are shown in Fig. 7. Curve *A*, obtained without a grid, has been included for comparison. We note that the slopes of the

curves are quite strongly affected and are considerably increased for small thicknesses.

However, the collimators lead to an appreciable reduction

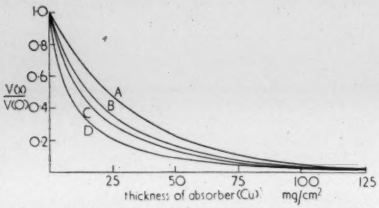


Fig. 7. The effect of collimation on the absorption curves

(^{204}Tl source; additional absorber $\simeq$ 10 mg/cm^2)

Curve *A*, without collimation; Curve *B*, with collimator *B*; Curve *C*, with collimator *C*; Curve *D*, with collimator *D*.

in β-ray intensity and hence in the signal-to-noise ratio $\alpha(0)$. This is shown in the table which gives the voltages developed

The effect of collimation on the output voltage

	Output voltage without absorber	*Output (coll.)* / *Output (uncoll.)*	$\alpha(0)$ *coll.* / $\alpha(0)$ *uncoll.*
Without collimation	7·58	—	—
With collimator *B*	2·09	0·275	0·525
With collimator *C*	1·37	0·181	0·425
With collimator *D*	0·396	0·052	0·228

across the ion chamber resistance with no absorbing material. The intensity reduction offsets the advantage gained from the increase in slope.

The function

$$\frac{\beta(x)}{\alpha(0)uncoll.} = \frac{\alpha(0)coll.}{\alpha(0)uncoll.} \cdot \frac{x(d/dx)V(x)/V(0)}{\sqrt{[V(x)/V(0)]}}$$

has been computed for each collimator and the resulting curves are shown in Fig. 8. Provided the accuracy of measurement is limited by output fluctuations, we note that it is consistently less when the grids are used than without collimation. This is true even for the thinnest materials, although curve *D* approaches curve *A* closely for small x.

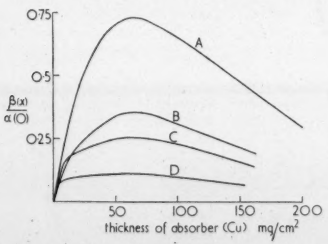


Fig. 8. The effect of collimation on the accuracy of measurement

(^{204}Tl source; additional absorber $\simeq$ 10 mg/cm^2)

Curve *A*, without collimation; Curve *B*, with collimator *B*; Curve *C*, with collimator *C*; Curve *D*, with collimator *D*.

While it has not been proved conclusively that considerable collimation or other geometrical arrangements cannot increase the accuracy, the indications are that the improvement is unlikely to be appreciable.

CONCLUSIONS

When the accuracy of thickness measurement is limited by the random fluctuations of the output, the percentage accuracy depends strongly on the thickness of material for any given β-ray source and has a maximum value. For an absorption curve of exponential form with absorption length a, the greatest accuracy is achieved when $x = 2a$. When the required percentage accuracy, the source strength and time constant are specified, there is a lower limit to the thickness which can be measured with any given source.

It appears unlikely that this limit can be appreciably reduced or the accuracy improved by the introduction of collimating grids.

ACKNOWLEDGEMENT

The writer wishes to thank W. A. Boyes for considerable assistance with both the experimental and numerical work.

REFERENCES

(1) SMITH. *Electronics*, **20**, No. 10, p. 106 (1947).
(2) BERTHOLD. *Z. Verein. Dtsch. Ingen.* [*VDI*], **95**, p. 207 (1953).
(3) ROSSI and STAUB. *Ionisation Chambers and Counters*, 1st ed., p. 25 (New York: McGraw-Hill Book Co., 1949).
(4) CURRAN, COCKCROFT and ANGUS. *Phil. Mag.*, **40**, p. 929 (1949).
(5) FANO. *Phys. Rev.*, **72**, p. 26 (1947).
(6) LANDAU. *J. Phys. U.S.S.R.*, **8**, p. 201 (1944).
(7) SYMON. Ph.D. Thesis, Harvard (1948).
(8) ROSSI. *High Energy Particles*, 1st ed., p. 32 (New York: Prentice-Hall, Inc., 1952; London: Constable and Co. Ltd., 1952).
(9) FANO. *Phys. Rev.*, **70**, p. 44 (1946).
(10) SCHIFF and EVANS. *Rev. Sci. Instrum.*, **7**, p. 456 (1936).
(11) RAINWATER and WU. *Nucleonics*, **2**, No. 1, p. 42 (1948).
(12) BURGESS. *Rev. Sci. Instrum.*, **20**, p. 964 (1949).
(13) SYKE. *Proc. Isotope Techniques Conference*, 1951, Vol. II, p. 144 (London: H.M. Stationery Office, 1952).

NOTES ADDED IN PROOF

Dr. Mandel has been good enough to let me see an advance proof of the above paper, and I am grateful to the Editor for allowing me to comment on it.

The author claims that in most applications of β-ray gauges statistical inaccuracies are large compared with those arising from other causes. Whilst statistical fluctuations can undoubtedly set the limit to β-ray gauge accuracy in some cases where high speed of response is essential, this is certainly not so in most industrial applications.

The factors which usually set the limit of attainable accuracy can be grouped broadly as follows:

(*a*) Variable absorbers in the measuring gap, additional to the strip one wants to measure. These include air in the gap and deposition of dust, oil, water (condensation) on source holder or ionization chamber windows.

(*b*) Changes in relative position of source, strip and detector; supporting members suffer small distortions due to temperature effects, or mechanical forces.

(*c*) Changes in the mass of free air contained in the ionization chamber due to diffusion and adsorption on walls.

(*d*) Accuracy limitations of the associated electrical measuring equipment.

In gauging very light materials, e.g. tissue paper, plastic foil, etc., below some 5 mg/cm^2 factors under heading (*a*) are significant; in case of heavier materials, say above 20 mg/cm^2, these factors are usually negligible. In effect, we are faced with a variable initial absorber; the variations actually encountered depend entirely on local conditions and may be expressed directly in milligrams per square centimetre.

Factors under headings (*b*), (*c*) and (*d*) cause some error in electrical output of the measuring equipment; it is logical and convenient to express this error as a fraction (or percentage) of the maximum output, V_0, obtained from one measuring head with initial absorber alone in the measuring gap. In a well-designed instrument under reasonably favourable factory conditions errors due to all factors under (*b*), (*c*) and (*d*) can usually be kept within $\pm 0{\cdot}2\%$ of maximum output, V_0.

What this error amounts to in terms of absorber thickness depends on $\Delta V/\Delta x$ whereby ΔV is the incremental change in instrument output due to an incremental change Δx in absorber thickness.

It is convenient to define a factor S:

$$S = \frac{\Delta V/V_0}{\Delta x/x}$$

whereby x is the thickness of the strip being measured. In practice the factor S is of the order of $0{\cdot}3$ in the central portion of the absorption curve (say 30% to 85% absorption) and decreases rapidly towards the ends. This means that in the central region 1% change in absorber thickness ($100\,\Delta x/x = 1$) will produce about $0{\cdot}3\%$ change in output ($100\,\Delta V/V_0 = 0{\cdot}3$). It will be readily seen from this, that $\pm 0{\cdot}2\%$ output error in the central region will give rise to about $\pm 0{\cdot}7\%$ thickness error.

To work in the central region when gauging very light materials we would require very low energy β sources. We are faced with two serious difficulties.

(1) Low energy β sources with long half life and high specific activity are not available at present.

(2) The minimum initial absorber for hermetically sealing the ionization chamber and for eliminating contamination hazard under factory conditions can be put at about 20 mg/cm^2. A β source of say 2 mg/cm^2 half value thickness, which would give optimum sensitivity and accuracy in measuring specimens of 1 to 3 mg/cm^2, would thus be impracticable (even if it could be obtained); the 20 mg/cm^2 initial absorber would absorb $99{\cdot}9\%$ of the radiation!

The collimator (Brit. Pat. 689857) enables us to use thallium 204 (half value thickness about 30 mg/cm^2) and still obtain a sufficiently high value for S down to absorber thicknesses of 1 or 2 mg/cm^2. A fascinating feature of the collimator is that the apparent absorption coefficient *between* the collimators, where the specimen is situated, is much higher than *outside* the collimators, where the windows of source holder and ionization chamber are situated. Any loss in radiation flux due to the collimator can usually be made good without increasing the stray radiation intensity, by using stronger source or larger ionization chamber.

A detailed analysis of β-gauge accuracy limitations and of collimator theory fall outside the scope of these notes. In conclusion it should be noted that:

(1) The collimator will reduce measuring errors due to factors (*b*), (*c*) and (*d*) when gauging light materials. This is the true purpose of using collimators.

(2) Errors due to statistical fluctuations of the radioactive disintegration process are not decreased just by adding collimators. Since, however, the limit of permissible maximum source strength is usually set by health hazard considerations, the collimator, which reduces stray radiation, makes it possible to use stronger sources and can thus indirectly also help to improve statistical accuracy.

(3) Statistical fluctuations set a limit to maximum useful speed of response at a specified accuracy, rather than to accuracy by itself.

(4) A well designed β-ray thickness gauge, as supplied by competent instrument manufacturers, correctly installed under reasonably favourable factory conditions will give overall accuracy of the order of $\pm 1\%$ over a large portion of its range.

G. SYKE

Baldwin Instrument Co.,
Dartford, Kent.

Mr. Syke challenges the relevance of some considerations on statistical inaccuracies in my paper and, in particular, my remarks on the use of grid collimators. The conclusion drawn in the paper was stated unambiguously, namely, "Provided the accuracy of measurement is limited by output fluctuations . . . it is consistently less when the grids are used than without collimation." Nothing was said about possible applications where the accuracy is not so limited.

Syke now dismisses this conclusion as largely irrelevant in practice with the remark—"Whilst statistical fluctuations can undoubtedly set the limit to β-ray gauge accuracy in some cases where high speed of response is essential, this is certainly not so in most industrial applications." He then proceeds to summarize—most ably—the factors, such as long-term drifts, etc., which in his view limit the accuracy most. Now both these cases are known in practice and it would clearly be futile to enter into a controversy on the question. Suffice it to say, however, that the random output fluctuations were almost invariably the limiting factor in the applications (particularly mill applications) with which the writer has been connected.

Since there still appears to be some confusion between the two cases it may be worth while to summarize the relevant facts. Accuracy in β-ray thickness measurement is limited by two factors, namely:

(*a*) random fluctuations of the output, as discussed in my paper;

(*b*) drifts and other long-term variations, as discussed by Syke (see above).

The inaccuracies arising under (*a*) depend strongly on the source strength and speed of response of the instrument. They are great for weak sources and high speeds and tend to zero as the response time is lengthened to infinity. The inaccuracies under (*b*) constitute the inevitable experimental errors associated with all measurements. They are independent of the response time (although not necessarily of the source strength), but are related in many ways to the conditions under which the instrument is working and to the designer's skill.

When a reasonably rapid response is required, the inaccuracies under (*a*) are predominant and these have been

treated in some detail. When the response time is a less important consideration, it is possible to slow down the instrument until the errors under (*a*) are rather smaller than those under (*b*), which are discussed by Syke (see above). The claim that these latter constitute the only important limiting factor is, however, quite unjustified.

The variation of measuring accuracy with absorber thickness when case (*a*) is predominant has already been evaluated, but it is not difficult to derive the corresponding results when case (*b*) is predominant. With the same notation, we define the accuracy of measurement $\beta(x)$ as:

$$\beta(x) = x/\Delta x \quad = -(1/\Delta V)x(d/dx)V(x) \qquad (1)$$

where Δx is the smallest reliably recognizable change in thickness.

We shall now suppose that the uncertainty ΔV, under subheading (*b*), consists partly of a constant voltage, independent of the source and due mainly to the electronics, which we shall express as a fraction ϵV_0 of the maximum V_0 and partly as a given fraction $\kappa V(x)$ of the output voltage $V(x)$, due to the other factors quoted by Syke. (Note that ϵ, so defined, varies inversely with V_0.) The signal-to-noise ratio $V(x)/\Delta V = V(x)/(\kappa V(x) + \epsilon V_0)$ will be denoted by $\alpha(x)$, so that $\alpha_0 = 1/(\kappa + \epsilon)$. Then, from equation (1)

$$\beta(x) = -\alpha_0(1 + \epsilon/\kappa)\frac{x(d/dx)V(x)/V_0}{V(x)/V_0 + \epsilon/\kappa} \qquad (2)$$

It is interesting to compare this result with equation (22) of my paper which applies when case (*a*) is predominant:

$$\beta(x) = -\alpha_0\frac{x(d/dx)V(x)/V_0}{\sqrt{[V(x)/V_0]}} \qquad (3)$$

There is a similarity of form, but equation (2) depends also on the ratio ϵ/κ. When dealing with an approximately exponential β-ray absorption spectrum of the type $V(x) = V_0\exp(-x/a)$, we may simplify the equations as follows:

$$\beta(x) = \alpha_0\frac{1 + \epsilon/\kappa}{\exp(-x/a) + \epsilon/\kappa}\frac{x}{a}\exp\left(-\frac{x}{a}\right) \qquad (4)$$

when case (*b*) is predominant and

$$\beta(x) = \alpha_0\frac{x}{a}\exp\left(\frac{x}{2a}\right) \qquad (5)$$

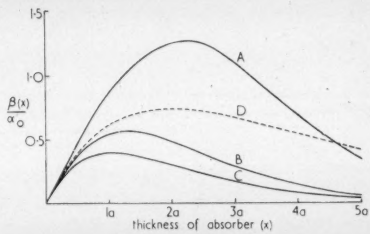
when case (*a*) is predominant.

The functions are compared in the figure for three values of the ratio ϵ/κ, namely, 0·1, 1·0 and 10. It should be borne in mind that $\alpha_0 = 1/(\kappa + \epsilon)$ for curves *A*, *B* and *C* but $\alpha_0 \simeq 1/\sqrt{(2N_0\tau)}$ for curve *D*.

We are now in a position to examine quantitatively the effect on the measuring accuracy of using the collimators first described by Syke.* Since these lead both to a decrease of β-ray intensity and to an initial decrease of the absorption length *a*, their use results in an initial compression of the *x*-scale for curves *A*, *B* and *C*. Provided α_0 does not decrease at the same time, this could be interpreted as an improvement in measuring accuracy for very thin materials.

Let us, however, examine the assumption that α_0 does not decrease. If κ is rather greater than ϵ, as for curve *A*, $\alpha_0 \simeq 1/\kappa$ and is approximately constant. The assumption is, therefore, valid and the collimators lead to an improvement in accuracy for thin materials.

Suppose now that κ is of the same order or less than ϵ. Since ϵ varies inversely with β-ray intensity, α_0 is now diminished due to the reduction of beam strength. There is, therefore, no longer an improvement in accuracy. In particular, if ϵ is rather greater than κ so that α_0 is practically proportional to V_0, the accuracy is much reduced by using collimators, for it has already been shown experimentally that the intensity reduction is more severe than the initial decrease of absorption length.



The variation of measuring accuracy with absorber thickness for an exponential absorption spectrum

Curve *A*, when case (*b*) is dominant with $\epsilon/\kappa = 0{\cdot}1$
Curve *B*, when case (*b*) is dominant with $\epsilon/\kappa = 1{\cdot}0$
Curve *C*, when case (*b*) is dominant with $\epsilon/\kappa = 10$
Curve *D*, when case (*a*) is dominant

We conclude, therefore, that the collimators lead to improved accuracy if and only if:

(i) statistical inaccuracies due to output fluctuations are small compared with other inaccuracies, and

(ii) electrical and other output drifts which are independent of source strength are small compared with those dependent on source strength.

If either or both of these conditions does not hold, the collimators lead to a worsening of measuring accuracy. It is hoped that the foregoing is an adequate reply to Syke's statement 1 above.

Statement 2 is somewhat euphemistic, since the grids have been shown to bring about an increase of statistical errors. The loss of accuracy can, in principle, be made good by increasing the source strength, but it is in any case the practice to use as strong a source as possible, when statistical errors are of importance.

Statement 3 is valid in applications in which the response time is a very secondary consideration and can be lengthened arbitrarily. However, when a high speed of response is specified, so that the statistical inaccuracies are dominant, the two are explicitly related and specification of either one imposes a limit on the other.

It is not intended to challenge statement 4. It may, nevertheless, be worth putting on record that the writer is familiar with potential mill applications of β-ray gauges in which it would be difficult to achieve 5% accuracy.

L. Mandel

Central Instrument Section,
Imperial Chemical Industries Ltd.,
Welwyn.

* Syke, G. *Proc. Isotopes Techniques Conf., Vol. II, Industrial and Allied Research Applications*, p. 144 (London: H.M. Stationery Office, 1952).

Evaporated carbon films for use in electron microscopy

By D. E. BRADLEY, Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks.

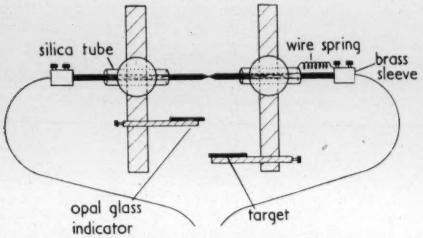
[*Paper received* 5 *August*, 1953]

A method has been developed for producing carbon films suitable for electron microscope specimen supports. Carbon is evaporated on to an extremely soluble substrate, which is dissolved away leaving very thin films.

It has been found possible to evaporate carbon to form uniform amorphous films suitable for electron microscope specimen supports. These films are exceptionally strong and can therefore be made extremely thin, so that, since they have a low atomic number, they are highly transparent to electrons.

The method of evaporation is to pass an alternating current of between 20 and 50 A through two pointed carbon rods in a vacuum chamber, with the points held lightly together so that they are not parted during the process. Intense local heating occurs in the region of contact.

The apparatus used is illustrated in the figure. The 0·5 cm diameter carbon arc rods are supported in 0·6 cm internal diameter silica tubing held in bosses screwed to two upright rods. To maintain a slight pressure between the pointed



Apparatus for the evaporation of carbon

ends, one electrode is fixed, and a small spring made from 0·01 in. tantalum or tungsten wire is fitted between the terminal on the other rod and a small hole blown in the silica tube. The leads are attached to the rods by short brass sleeves, fitted with grub screws.

The flat piece of white glazed porcelain or opal glass with a drop of Apiezon "B" oil on it, mounted on a movable platform, helps in the estimation of the film thickness during evaporation. The condensed carbon is clearly visible on the porcelain but not on the oil drop, which shows up in sharp contrast. When this is just visible, the thickness of the film on the indicator is about 100 Å (the actual thickness was measured by the optical interference method of Tolansky*). A suitable current is obtained from a 24 V high current transformer. A typical evaporation is carried out with the indicator placed at 8 cm from the source and the target slide at 13 cm. Assuming that the evaporation of carbon obeys the inverse square law, the thickness of the film on the target should be about 40 Å in this case. However, pilot experiments have shown that the thicknesses of the carbon deposits obtained in this way fall off more rapidly than predicted by the inverse square law, and so the calculated thickness will be greater than the real thickness.

The film is evaporated on to a substrate which can easily be stripped from the glass slide. A suitable substrate has been found to be Bedacryl 122 X, supplied in a 40% solution in Xylene or in solid form by Imperial Chemical Industries Ltd. Boron oxide or glycerol can also be used and the film mounted from a water surface, but the thinner films break with these substrates.

To coat the target slide, the Bedacryl is first diluted with redistilled benzene to a strength of 6–8% weight/volume. A quantity is then poured over a clean microscope slide and allowed to drain off. The film of resin dries in a few seconds under a lamp.

After evaporation, the film is scored into small squares and floated on to a water surface. Stripping is made easier by breathing heavily on to the film, as with formvar, and some teasing at the edges may be necessary, otherwise Bedacryl strips very easily.

A specimen grid is now slightly bent in such a way that it lies on a flat topped peg supported at the edges only. Alternatively, a peg with a convex top can be used and the bending can be omitted. The bending of the grid or the use of a round-topped peg prevents contact of the film with the surface of the peg over a large area where the film would otherwise subsequently break.

The floating squares are picked up with a lifter consisting of a bent metal strip with a $\frac{3}{16}$ in. diameter hole in it. They are then inverted over the grids mounted on the pegs in the order: grid, carbon, Bedacryl. If the hole in the lifter is too small, the resin film may break.

At this stage, it has been found necessary to bring the carbon into intimate contact with the grid surface, otherwise the thin film will float off in subsequent washing. This is brought about by rendering the Bedacryl tacky, and then drying it. Thus, four or five drops of methylated spirits are allowed to fall on the specimen and surplus liquid is removed with the edge of a filter paper. When the surface only of the resin film is dry, the grid is lifted in a pair of forceps. If the substrate is thin enough, the grid will come free quite easily and drying is completed in a few seconds. To wash off the substrate, the forceps are held at about 45° below the nozzle of a burette containing a mixture of 50% ether and 50% acetone. About 1·5 c.c. of the mixture is allowed to run slowly down the forceps and over the grid. Surplus liquid should be removed from between the forceps with filter paper or the specimen will not dry quickly. A thin carbon film should now cover the grid.

For thicker films the grid need not be removed from the peg, but the drying of the methylated spirits should be complete. The grid may be washed while on the peg by allowing the solvent mixture to flow over but not drop on to the specimen. Drying should be carried out under a lamp.

This method of mounting has been found satisfactory for the thinnest films, though the normal 200 mesh to the inch grids are only suitable for rather thicker films. Grids made with 0·0015 in. circular holes were found more suitable in supporting very thin films. These stood up to heavy electron bombardment and supported specimens of reasonable size. To mount specimens from water suspension, grids should not be placed on the normal type of peg, but on short lengths of

* TOLANSKY, S. *Multiple Beam Interferometry of Surfaces and Films* (London: Oxford University Press, 1948).

thin-walled $\frac{1}{8}$ in. brass tube fitted with caps. It is more satisfactory to deposit the specimen before removing the substrate, provided it is not affected by the solvents.

The above method of mounting has been applied to silica films made by evaporating silicon monoxide. Again it has been found possible to mount very thin films, although they are not as strong as those made with carbon, nor are they as transparent to electrons.

It has been observed that the structure of carbon films is much finer than the ordinary collodion films and also than beryllium films. In some cases there appeared to be no structure at all after the film had been shadowed.

ACKNOWLEDGEMENT

The author wishes to thank Dr. T. E. Allibone for permission to publish this note.

High vacuum applications of polythene

By J. F. Duncan, M.A., D.Phil.,* and D. T. Warren, B.Sc., Atomic Energy Research Establishment, Harwell, Berks.

[*Paper first received* 19 *June, and in final form* 26 *September*, 1953]

The use of polythene in high vacuum work is discussed. The fabrication of polythene vessels, valves, taps, fine control valves, manometers, etc., is described, together with the use of polythene in jointing glass, metal, alumina and other ceramics to high vacuum lines.

For the construction of high vacuum apparatus, it is conventional to use one of two techniques. One may build in glass, with the advantages of visibility, a very good vacuum, and extremely clean conditions, but with the attendant disadvantages of fragility, inability to make machined surfaces and limitations on the temperature range of operation. Alternatively, one may build in metal, when one obtains very robust equipment, which can be machined with extreme accuracy if necessary; but one cannot usually obtain such a high vacuum as with glass, and the risk of contamination with oils, vapours, solder flux, etc., is always present. For the majority of high vacuum investigations, one of these techniques, or a combination of them, is suitable, but sometimes chemical considerations preclude both of them. Glass, for instance, strongly absorbs water vapour, and metal is not very much better. More serious is the possibility of reaction with materials such as hydrogen fluoride, fluorine, chlorine and its oxides and other reactive gases. Alternatively, one may wish to develop high vacuum apparatus which is flexible, capable of being machined, and which is not likely to introduce toxic substances (in the pharmacological sense) or materials which poison chemical reactions. For all these applications, and probably for others too, polythene is a possible constructional material. The techniques described below were, in fact, developed to overcome difficulties in the use of hydrogen fluoride under high vacuum conditions.

HIGH VACUUM PROPERTIES OF POLYTHENE

Polythene tubing of $\frac{1}{8}$ in. wall thickness and not more than $\frac{1}{2}$ in. diameter will withstand complete evacuation without collapsing. Incidentally, it will also withstand internal pressures of up to 200 lb/in.^{2} For vessels of 3 in. diameter, wall thicknesses up to $\frac{1}{8}$ in. may be necessary to prevent collapse. As with other material, pumping in a high vacuum for several days is necessary to remove all adsorbed gases, but pressures of the order of 10^{-6} mm of mercury can easily be obtained when the wall surfaces are clean. Polythene is known to be slightly permeable to gases and water vapour, but with reasonably thick walls, this is no difficulty. There is some tendency for dissolved gases to be slowly liberated from the polythene wall, but this is also not serious in a carefully pumped apparatus. Pressures of 10^{-4} mm of mercury can be held overnight when the apparatus is clean. To assist removal of occluded gases a low potential Tesla discharge coil is usually used. A naked flame is to be avoided, as the melting point of polythene is only about 110° C.

If an apparatus having some parts made of polythene has been properly cleaned by pumping and is exposed to atmospheric air for a few minutes and subsequently re-evacuated, it is found that the glass and metal parts are invariably heavily contaminated by adsorbed gases and require pumping for some hours. The polythene parts of the apparatus are, however, usually fairly clean, and when these alone have been exposed to the atmosphere, pumping for about half an hour is usually adequate to remove all the adsorbed gases.

FABRICATION OF POLYTHENE APPARATUS

Polythene can be joined to itself, to glass and to metal by welding with a jet of hot nitrogen. This technique, which is well established and used industrially, is adequate for making joints which can be pumped down to pressures of 10^{-6} mm of mercury. It is even possible to join sheets of 0·005 in. thick polythene to polythene tubing by this technique. Polythene can also be machined, and this has advantage in the fabrication of valves, and apparatus containing moving parts. The accuracy of machining is limited in use by the tendency of the polythene to flow, but this is not serious in apparatus not subjected to excessive strain or compression. In cases where excessive pressure is anticipated it is often better to use a harder plastic material such as polytetrafluoroethylene.

Polythene is also available in powder form, suitable for coating metal and glass components. The component is heated to about 200° C and plunged into the powder, which then sticks to it. Heating to about 120° C produces an adherent uniformly thick coating, which will protect the component from chemical attack, etc.

Tubing may often be straightened, or bent, by inserting a strong coil spring as a support, immersing the tubing in a bath of hot water for 15 min, and fixing the tubing in the desired shape. After cooling the tubing retains its shape.

HIGH VACUUM COMPONENTS MADE OF POLYTHENE

All the components described below have been tested under high vacuum conditions and used with corrosive gases. The designs described are those in current use only. Other less satisfactory designs were employed earlier.

Vessels. Polythene vessels up to 6 in. diameter may be made by sealing the ends of a $\frac{1}{2}$ in. thick polythene tube with disks. For smaller vessels, it is preferable to use drilled rod and to seal round a lapped joint on the cylindrical face (see Fig. 1).

* Now at Chemistry Department, University of Melbourne.

Valves. A diagram of a polythene valve is shown in Fig. 2. The top end of the chamber, *A*, is about $\frac{1}{32}$ in. thick and is sufficiently flexible to allow the stud, *B*, to seat on the end of the tubing, *C*, when compressed by the brass plunger, *D*. The metal spring, *E*, which pushes the stud back when the valve is open, is polythene-coated by the technique described

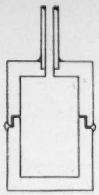


Fig. 1. Construction of small polythene vessels, by use of lapped joint and seal

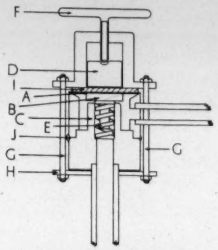


Fig. 2. Polythene valve

A, flexible diaphragm; *B*, stud; *C*, valve seating; *D*, brass plunger; *E*, polythene-coated metal spring; *F*, brass key; *G*, brass clamping rods; *H*, brass retaining ring; *I*, neoprene washer to protect polythene from abrasion by plunger; *J*, welding junction.

above. The whole interior of the valve is thus made of polythene and no metal parts are in contact with the gas. The brass superstructure supporting the key, *F*, and plunger, *D*, is held to the polythene body by six brass rods, *G*, and a retaining ring, *H*. The whole valve is welded at the junction, *J*, after complete assembly.

Taps. In principle, one can also fabricate taps of conventional type in polythene, but they are not suitable for high vacuum work even with good machined surfaces. Owing to the softness of the polythene, the barrel of the tap tends to deform, with consequent leakage; but such taps are satisfactory for use with liquids.

Fine control valves. A diagram of a fine control valve is shown in Fig. 3. This valve is of conventional type, except

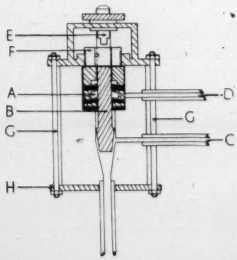


Fig. 3. Fine control valve

A, Wilson seal; *B*, polytetrafluoroethylene spindle; *C*, gas flow; *D*, lead for pumping out Wilson seal; *E*, brass thread, locked to spindle *B*; *F*, brass screw; *G*, brass clamping rods; *H*, brass retaining ring.

for two particulars. First, a double Wilson seal, *A*, is used to prevent entry to the atmosphere into the chamber of the valve. These seals were made with neoprene rubber, lubricated with polythene grease (see below), although conventional greases are satisfactory if the gas does not chemically attack them.

Secondly, a polytetrafluoroethylene spindle, *B*, is used to control the flow. For reproducibility, the spindle should be of hard material, and polythene is not suitable since it tends to deform. Polytetrafluoroethylene cannot be easily bonded to any material, and therefore it was necessary to make the entire spindle of it. This incidentally results in a better seal at the neoprene ring than could be obtained with a polythene spindle, although polythene would be quite satisfactory here.

Manometers. Since it was required to measure the pressure of corrosive gases, it was essential to use only non-reactive material. The pressure was measured by balancing the gas

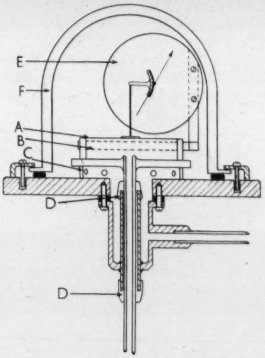


Fig. 4. Polythene manometer

A, polythene diaphragm; *B*, polythene tube; *C*, polythene support; *D*, polythene–metal joints; *E*, scale indicator; *F*, bell jar.

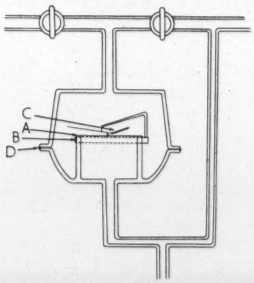


Fig. 5. Glass manometer with polythene diaphragm

A, polythene diaphragm; *B*, clamping ring; *C*, quartz movement; *D*, seal of Apiezon W wax.

pressure across a polythene membrane against an air pressure which was measured on a differential oil manometer. Two forms of the apparatus are shown in Figs. 4 and 5. The manometer shown in Fig. 5 is identical in principle with that shown in Fig. 4, but, except for the polythene membrane, it is made entirely of glass. It has been used for measuring fluorine pressures over a period of 12 months without any sign of deterioration of the polythene membrane other than mechanical wear occasioned by the quartz fibre movement.

In Fig. 4 the diaphragm, *A*, about 0·02 in. thick, is welded

on to the polythene tube, *B*, about 5 in. diameter. To obtain a uniform tension on the membrane, it is welded first on one side, and then on the opposite side of the tube, while still being held under tension. It is next welded in the transverse direction in the same way, and this is repeated until the diaphragm is sealed all round. A circular wedge is also often inserted in the tubing after welding in order to maintain the tension, as such large diameter polythene tubes often deform to an elliptical shape if left for long periods. After insertion of the wedge, the lower end of the tube which supports the membrane is welded on. The whole is supported on the polythene neck piece, *C*, for stability. An internally sealed metal (or glass) tube is next sealed at both ends, *D*, by welded polythene to a piece of polythene tube, which passes through it. The top end of the tube is connected to the membrane chamber, and the bottom end to the pressure system. Thus the gas is not exposed to any material other than polythene.

A lightly pivoted Bourdon manometer pinion indicator, *E*, is allowed to rest on the top of the membrane, *A*, movement of which is indicated by the needle. On the original manometers designed by us, a quartz fibre movement was used. This is quite satisfactory but rather fragile. The modification later introduced by Waldron and Wilkinson makes the instrument more robust without any significant loss of sensitivity. The whole is covered with a bell jar, *F*, sealed with a neoprene rubber ring.

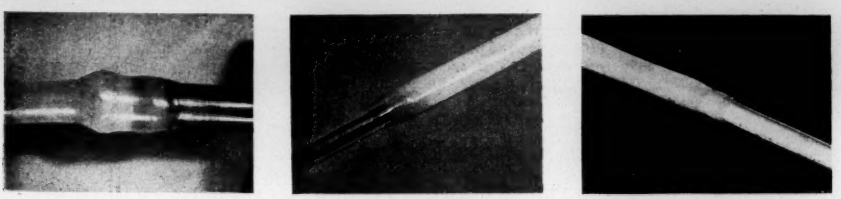
In the all-glass manometer (Fig. 5) the polythene membrane is sealed to the glass with a thick high vacuum grease and held by a clamping ring. The lead to the air pressure side of the system was through the bell jar in this case. This simplified the construction of the lower part of the manometer, but sacrificed the advantage of being able to repair the movement without dismantling the apparatus, as is possible with the manometer shown in Fig. 4. The quartz fibre movement (observed by means of a telescope) is clamped by another ring on top of the ring which holds the membrane in position.

Both manometers are used as null deflexion instruments, the absolute pressure being obtained from a differential oil manometer connected between the air chamber and a good vacuum. By this technique the sensitivity is limited by that of the auxiliary pressure-measuring device, which in this case is sensitive to about 5×10^{-2} mm of mercury. If the instrument were used as a direct reading instrument, the sensitivity could be improved by at least tenfold. The polythene membrane will stand pressure differences of 100 mm of mercury without rupture, but to avoid zero errors due to stretching of the membrane, it is better to keep the maximum pressure difference across the membrane below 10 mm of mercury.

Liquid air traps. Polythene will not withstand liquid air temperatures unless it is unstrained, and is not reliable for use in liquid air traps unless precautions are taken to avoid steep temperature gradients. If such an application is unavoidable, it is desirable to use polythene of substantial wall thickness, so that strain is not introduced due to deformation of the walls on account of the high vacuum. Immersion of welded joints in the liquid air should also be avoided since it is very difficult to get them strain free. The best technique is, therefore, to make the trap by machining out of a solid piece of polythene, and to connect polythene leads externally to it by welding.

If the trap should contain any strained parts, it will shatter on immersion in liquid air, but it has been reported[1] that this does not occur if the polythene is enclosed in a copper block. Apart from the risk of shattering, however, the low thermal conductivity of polythene makes it a poor constructional material for liquid air traps.

Bonding of polythene to other material. Joints between polythene and metal, glass, alumina and other ceramic materials may be made by heating the end of the metal or glass tube above the softening point of polythene and inserting it into the end of the polythene tube. The assembly is again gently warmed and as the polythene begins to flow it is pressed out by means of a flat cold metal tool, to give good vacuum-tight adhesion. Such joints can be pumped down to 10^{-6} mm of mercury. Polythene joints of this kind are often convenient for fabricating glass–metal, or glass–ceramic joints in high vacuum apparatus. Fig. 6 shows various joints of this kind.



(*a*) Copper–polythene joint (*b*) Glass–polythene joint (*c*) Alumina–polythene joint

Fig. 6. Bonding of polythene to other materials

Polythene grease. A small quantity of an experimental sample of polythene grease was kindly given us by I.C.I. Ltd., Plastics Division, Welwyn. This grease (I.P. drop point 127° F) has been used on the Wilson seals of the valve, and on the moving parts of a rotary reaction vessel.[2] It is, we understand, essentially a mixture of low molecular weight polymers. As it has a pronounced smell, it clearly has a high vapour pressure. After pumping down for a few hours, however, it was found that the pressure could be held below 10^{-3} mm of mercury for extended periods, presumably because the lower polymers had been pumped off. We have found it quite satisfactory for this kind of application.

VACUUM TIGHTNESS AND VACUUM PROVING

After a little practice, one can in almost all cases make polythene joints which are quite vacuum tight without further treatment. If a leak should be present, it may be hunted down by conventional methods using a hydrogen, helium or carbon dioxide probe. One can also probe with hot nitrogen or even (with care) a naked flame, using a Pirani or ionization gauge as indicator. The latter requires some skill,

as the polythene has a low melting point, but with practice it is quite a possible technique. When the leak has been found it is preferable to remake the joint entirely, but if the leak is small and it is essential not to lose the vacuum, the joint can, with practice, be mended in position with the apparatus still evacuated. This is done, often without exact location, by a welding technique using fresh polythene which is melted and pressed into the affected area. The finger may be used for this purpose and is a good means of ensuring that the polythene does not become hot enough to rupture; but a well-cooled flat metal tool is satisfactory and gives a neater appearance. Several leaks were mended by these techniques when polythene was first used for high vacuum work, but latterly we have been able to get almost all our joints vacuum-tight on first fabrication.

CONCLUSIONS

The apparatus described represents only a first step in the use of polythene for high vacuum work. There are still innumerable possibilities to be explored. Polythene has, of course, several disadvantages, such as its opaqueness and its tendency to flow, but these are compensated by some unique properties. We are convinced that it will assist the solution of many of the difficulties associated with the use of glass and metal apparatus under high vacua.

ACKNOWLEDGEMENTS

We are grateful to Dr. W. Fergusson of I.C.I. Ltd., Plastics Division, Welwyn, for supplying the polythene grease and for helpful discussions, and to Mr. P. Brown for assistance in developing the all-glass membrane manometer. The director of the Atomic Energy Research Establishment has given permission for this paper to be published.

REFERENCES

(1) WALDRON, J. C. H., and WILKINSON, K. L. Private communication.
(2) DUNCAN, J. F., and WARREN, D. T. *J. Sci. Instrum.*, **30**, p. 462 (1953).

The effect of the finite light sources of measuring instruments on the determination of electron diffraction ring radii

By G. D. ARCHARD, A.INST.P., *The University, Reading

[*Paper first received* 17 *April and in final form* 6 *July*, 1953]

The presence of diffuse background is known to displace the intensity peaks of electron diffraction rings. The form of the necessary correction is modified by the finite dimensions of the slits which constitute the effective light sources used in the measuring instruments. It is shown that a simple form of correction is restored when the position-determining and intensity-measuring instruments have the same sized slits.

Rymer and Butler[1] have shown that the presence of diffuse background on electron diffraction patterns causes a displacement of intensity peaks; the displacement is in the direction of greater background (Fig. 1). Thus the combined intensity distribution of a line of Gaussian form superimposed on a linearly decreasing background is:

$$I = m(r - r_0) + A \exp[-(r - r_0)^2/2\sigma^2] \qquad (1)$$

where r represents distance measured across the line and σ is the standard deviation. In electron diffraction, background generally decreases with increasing radius; hence ring radii will appear to be diminished.

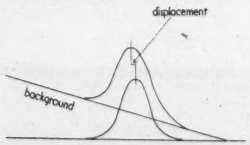


Fig. 1. Effect of diffuse background on the peak of an electron diffraction line

In order to locate the maximum intensity of the combined distribution, equation (1) must be differentiated with respect to r, so that:

$$m - A(r - r_0) \exp[-(r - r_0)^2/2\sigma^2]/\sigma^2 = 0 \qquad (2)$$

Provided that $r - r_0$ is small compared with σ, this may be written:

$$r - r_0 = m\sigma^2/A \qquad (3)$$

Expression (3) represents the displacement of the peak due to the background; it is in effect the expression given by Rymer and Butler,* where:

$$\sigma = W \times 0{\cdot}425 \qquad (4)$$

and W is the line "width" defined as the distance between the two positions at which the intensity is half that of the peak.

Whereas it is possible to determine the positions of the apparent peaks of electron diffraction rings by means of a travelling microscope, much greater precision can be obtained by means of a scanning microphotometer such as that described by Rymer and Halliday[2].

The measurement of intensity distribution required for the correction term (3) may be obtained with a second microphotometer; in a typical case, light from a slit source passes through the diffraction pattern and falls upon a thermopile which actuates a recording galvanometer. Such instruments incorporate slits of finite size, and this fact has the effect of modifying the form of background correction. The modifications due to the finite slit of the position-determining instrument will be considered first, then those due to the finite slit of the intensity-measuring instrument. Finally the effects will be combined.

* Now at Associated Electrical Industries Ltd., Aldermaston, Berkshire.

* Rymer and Butler showed that for any other intensity distribution likely to occur in electron diffraction the background correction based on a Gaussian distribution would merely require multiplication by a constant of order unity.

EFFECT OF THE FINITE SLIT OF THE POSITION-DETERMINING INSTRUMENT

Fig. 2 shows the trace of the slit passing over the contour map of a diffraction ring. The length y of a particular intensity contour inside the trace of the slit varies as shown in Fig. 3, wherein x represents the distance of penetration of the contour into the slit trace.

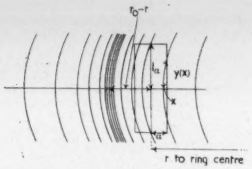


Fig. 2. Trace of a finite slit moving over the contour map of an electron diffraction line

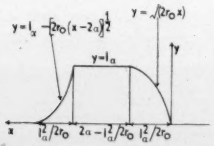


Fig. 3. Length y of an intensity contour falling within the slit trace as the slit moves across a diffraction line

The quantity $F(r)$ of light stopped by the diffraction ring segment inside the trace of the slit (and controlling the output of the microphotometer) may be written:[3]

$$F(r) = \int_0^{2\alpha+l_\alpha^2/2r} y(x) f(r - \alpha + x)dx \tag{5}$$

where r is the distance of the slit centre from the centre of the ring system, r_0 is the distance of the ring peak from the centre of the ring system, $2l_\alpha$, 2α, are the length and width of the slit trace, and where $f(\rho)$ is the function representing the variation of blackening of the photograph with radius ρ so that for a line of Gaussian form with a linear background:

$$f(\rho) = m(\rho - r_0) + A \exp\left[-(\rho - r_0)^2/2\sigma^2\right] \tag{6}$$

Replacing r by r_0 in the upper limit of the intensity integral, as $r - r_0 \ll r$, expression (5) may be expanded into:

$$\begin{aligned} F(r) = &\int_{l_\alpha^2/2r_0}^{2\alpha} l_\alpha \left\{ m(r - \alpha + x - r_0) \right. \\ &\left. + A \exp\left[-(r - \alpha + x - r_0)^2/2\sigma^2\right] \right\} dx \\ &+ \int_0^{l_\alpha^2/2r_0} \sqrt{(2r_0 x)} \left\{ m(r - \alpha + x - r_0) \right. \\ &\left. + A \exp\left[-(r - \alpha + x - r_0)^2/2\sigma^2\right] \right\} dx \\ &+ \int_{2\alpha}^{2\alpha+(l_\alpha^2/2r_0)} \left\{ l_\alpha - [2r_0(x - 2\alpha)]^{\frac{1}{2}} \right\} \left\{ m[r - \alpha + x - r_0] \right. \\ &\left. + A \exp\left[-(r - \alpha + x - r_0)^2/2\sigma^2\right] \right\} dx \end{aligned} \tag{7}$$

To locate the maximum intensity this must be differentiated with respect to r.

If the slit has negligible *length* only the first of the three integrals (7) will be of importance. It is easy to show that the correction expression corresponding to expression (3) will now be:

$$X = r - r_0 = (m/A)\sigma^2 \exp(\alpha^2/2\sigma^2) \tag{8}$$

provided that $X \ll \sigma$.

If the slit length may *not* be neglected, the second and third integrals in (7) must be considered. These are small compared with the first, and may therefore reasonably be written in the approximate form:

$$\begin{aligned} &\int_0^{l_\alpha^2/2r_0} \sqrt{(2r_0 x)} \left\{ m(X - \alpha + x) \right. \\ &\qquad \left. + A \exp(-\alpha^2/2\sigma^2)\left[1 + \alpha(X + x)/\sigma^2\right] \right\} dx \\ &+ \int_{2\alpha}^{2\alpha+(l_\alpha^2/2r_0)} \left\{ l_\alpha - [2r_0(x - 2\alpha)]^{\frac{1}{2}} \right\} \left\{ m(X - \alpha + x) \right. \\ &\qquad \left. + A \exp(-\alpha^2/2\sigma^2)\left[1 - \alpha(X + x - 2\alpha)/\sigma^2\right] \right\} dx \end{aligned} \tag{9}$$

If this expression be added to the first integral of (7) and the whole be differentiated with respect to $X (= r - r_0)$, the correction expression will become:

$$X = (m/A)\sigma^2 \exp(\alpha^2/2\sigma^2) - l_\alpha^2/6r_0 \tag{10}$$

which is the sum of the expression (8) and the "slit length" correction more simply obtained by Rymer and Butler.[1] The "slit length" term does not contain the background slope m and so may not properly be called a "background" correction; in practice, it is usually negligible.

EFFECT OF THE FINITE SLIT OF THE INTENSITY-MEASURING INSTRUMENT

The apparent intensity distribution has been given in expression (7), but, as the position-determining and intensity-measuring instruments are in general distinct, the slit dimensions α and l_α will be replaced by β and l_β.

With the same approximations as before, the expression for the peak intensity may be shown to be:

$$\begin{aligned} 4\beta l_\beta [&m(X + l_\beta^2/6r_0) - A \exp(-\beta^2/2\sigma^2) \\ &(Xl_\beta^2/6r_0\sigma^2 + l_\beta^4/40r_0^2\sigma^2) \\ &+ A(\sigma/\beta)\sqrt{(\pi/2)} \operatorname{erf}(\beta/\sigma\sqrt{2})] \end{aligned} \tag{11}$$

as compared with $4\beta l_\beta A$ which would have been observed if the whole slit had transmitted the same intensity as the peak.

For practical purposes the error function part of expression (11) predominates, so that very nearly:

$$A' = A(\sigma/\beta)\sqrt{(\pi/2)} \operatorname{erf}(\beta/\sigma\sqrt{2}) \tag{12}$$

where A' is the apparent peak intensity.

The relation of real to apparent line *width* may be obtained sufficiently accurately by considering a slit of negligible *length*. It is easy to show that the intensity (relative to the peak) recorded when the centre of the slit is distant z from the centre of the line (Fig. 4) is:

$$\frac{\operatorname{erf}[(z + \beta)/\sigma\sqrt{2}] - \operatorname{erf}[(z - \beta)/\sigma\sqrt{2}]}{2 \operatorname{erf}(\beta/\sigma\sqrt{2})} \tag{13}$$

It is desired to determine the value of $z = z_{\frac{1}{2}}$ for which

this quantity equals $\frac{1}{2}$. Then the apparent standard deviation σ' will be:

$$\sigma' = 0{\cdot}85\,.\,z_{\frac{1}{2}} \qquad (14)$$

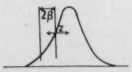


Fig. 4. To illustrate the intensity recorded by a microphotometer with finite slit at a distance z from the intensity peak of a diffraction line

The determination is best performed graphically; typical results are given in the table.

σ/β	σ'/σ
10·0	1·004
5·0	1·009
2·5	1·026
1·67	1·066
1·25	1·116
0·83	1·258

COMBINED EFFECTS OF THE FINITE SLITS OF THE TWO INSTRUMENTS

By combining expressions (10) and (12), the following equation is obtained:

$$X = (m/A')\sigma^2 \exp(\alpha^2/2\sigma^2)(\sigma/\beta)\sqrt{(\pi/2)}\,\mathrm{erf}\,(\beta/\sigma\sqrt{2}) \qquad (15)$$

(apart from the slit length correction omitted for reasons previously given).

Had the finite slit dimensions not been taken into account, an apparent background correction X' would have been obtained, in the form:

$$X' = (m/A')\sigma'^2 \text{ [compare equation (3)]} \qquad (16)$$

This may be related to the true background correction as follows:

$$X/X' = [\sigma^2 \exp(\alpha^2/2\sigma^2)/\sigma'^2](\sigma/\beta)\sqrt{(\pi/2)}\,\mathrm{erf}\,(\beta/\sigma\sqrt{2}) \qquad (17)$$

in which σ'/σ is a function of σ/β for which some values have been recorded in the table.

In Fig. 5 the values of X/X' are plotted against α/σ for various values of β/σ. This shows that if the slit of the position-determining instrument is perfectly narrow ($\alpha = 0$), but that of the intensity-measuring instrument is of the same order of width as the diffraction line, then the apparent background correction X' will be *double* the true correction X. On the other hand, if the slit of the intensity-measuring instrument is perfectly narrow ($\beta = 0$), but that of the position-determining instrument is of the same order of width as the diffraction line, then the apparent background correction X' will be *half* the true correction X. In practice it is not always possible to make $\alpha \ll \sigma$ so that errors of the order indicated are not inconceivable.

Consider a diffraction ring of radius 2 cm. If there exists a fairly steep background slope in this region, a correction of 10 μ would not be unusual. If, however, no allowance is made for the effect of finite slits, this may be only half the true correction. Consequently an error of 1 in 2000 would accrue in any lattice constant deduced from this ring.

Further considerations of Fig. 5 reveals a useful property. It will be observed that on the line $X/X' = 1$ the values of α/σ and β/σ are very nearly equal for a range of slit widths varying from zero to the width of the line. Thus, provided that the slits of the position-determining and intensity-measuring instruments are of equal size, the simple expression (16) may be used for the background correction, A' and σ' being the peak intensity and standard deviation obtained from the intensity record without correction for finite slit width.

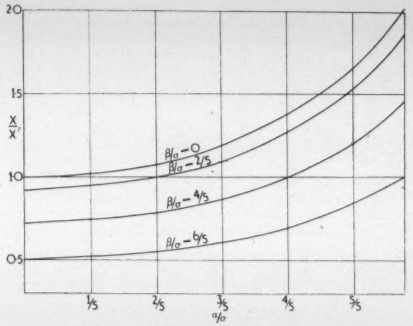


Fig. 5. Ratio of real to apparent background correction for various ratios of slit widths to diffraction line widths

α = slit half-width of position-determining instrument;
β = slit half-width of intensity-measuring instrument;
σ = standard deviation of (Gaussian) diffraction line.

This condition would be fulfilled automatically if the same instrument were used for both purposes; it would then be reasonable, for any one diffraction pattern, to use a slit of the same order of width as that of the narrowest line.

CONCLUSION

It has been shown that under circumstances which may be approximated in practice the correction of electron diffraction ring radii for diffuse background may be halved or doubled on account of the finite slits of the measuring instruments. This may result in errors in the lattice constants deduced from the rings, rising in extreme cases to 1 part in 2000. An expression has been given for the correction of this effect, but it has further been shown that no modification to the simple background correction is required if the slits constituting the effective light sources of the position-determining and intensity-measuring instruments are of the same size.

ACKNOWLEDGEMENTS

I wish to thank Professor R. W. Ditchburn for laboratory facilities, Dr. T. B. Rymer for helpful discussions, and the Department of Scientific and Industrial Research for a grant which was operative at the time of the investigation.

REFERENCES

(1) RYMER, T. B., and BUTLER, C. C. *Phil. Mag.*, **36**, p. 515 (1945).
(2) RYMER, T. B., and HALLIDAY, J. S. *J. Sci. Instrum.*, **27**, p. 50 (1950).
(3) BUTLER, C. C. Ph.D. Thesis, University of Reading, pp. 54–9 (1946).

A photometric method of determining configuration factors

By D. I. Lawson, M.Sc., M.I.E.E., F.Inst.P., and D. Hird, B.Sc., Fire Research Station, Boreham Wood, Herts.

[*Paper first received* 14 *May, and in final form* 19 *June*, 1953]

A photometric method is described which enables a speedy exploration of configuration factors to be made. The method can also be used to solve the reverse problem of finding the positions with respect to a radiator at which there is a given configuration factor. Whereas the method can be applied more easily to two-dimensional sources of radiation, it can also be applied to many three-dimensional radiators.

It is frequently necessary to calculate the radiation falling on to a receiving surface in the neighbourhood of a radiator maintained at some given temperature. Calculations of this type occur in furnace design, illumination engineering, radiant heating and studies on the growth of fire. These are often very tedious to carry out and the purpose of the present publication is to outline a simple method whereby the calculations may be performed mechanically with an accuracy of a few per cent.

The intensity of radiation at the surface of a "black body" at a temperature $T°$ K is given by

$$I = \sigma T^4 \qquad (1)$$

where σ is the Stefan–Boltzmann constant ($1{\cdot}37 \times 10^{-12}$ cal per cm^2 per sec per °C^4). The intensity of radiation falling on a receiver at any distance from the radiator not only depends on the temperature of the radiating area, but is also a function of the size of the radiator, and the distance and orientation of the receiving element with respect to it. Thus the intensity of radiation at any point with respect to the radiator can be expressed as

$$I_1 = \phi\, \sigma T^4$$

where ϕ is the configuration factor which may vary between 0 and 1, which takes into account the geometry of the problem. When the receiver is close to the radiator the configuration factor is unity giving $I = \sigma T^4$, and when the receiver is at a great distance from the radiator ϕ will have a value tending to zero.

The integral for the configuration factor (see Appendix) is given in most books on heat transfer. The problem does not lend itself well to mathematical treatment in any but the simple cases, examples of which have been worked out mathematically by Barker and Kinoshita,[1] McGuire[2] and some other workers. The numerical integration can be replaced by a graphical procedure first mentioned by Herman,[3] by an optical projection method[4] or by a mechanical integrator.[5] All these methods depend on the comparison of two areas to give the configuration factor, and the whole procedure must be repeated for each determination of ϕ at different positions with respect to the same radiator. By none of these methods is it possible to do the reverse problem, i.e. to find at what position with respect to a radiator there is a given configuration factor. The method described in this paper can be used to solve this problem and gives a speedy method of making a complete exploration of the configuration factor with respect to a radiator. In its present form the apparatus can only solve problems in which the radiating surface is two-dimensional or can be approximated to a two-dimensional equivalent surface.

Principles of the Method and Apparatus

Since heat and light are propagated according to the same laws, then if a surface radiating heat is replaced by one radiating light under similar conditions, a photocell detector placed close to the surface of the radiator will give a reading representing σT^4 in the heat problem [equation (1)]. When placed in the position similar to that of the receiving element it will give a reading reduced by the factor ϕ, the configuration factor. Since the configuration factor is a pure number, the linear dimensions of the radiator can be scaled by any factor providing the distance of the receiver is scaled by the same factor.

An apparatus exploiting this analogy is shown in Fig. 1. A model of the radiating area cut out of black photographic

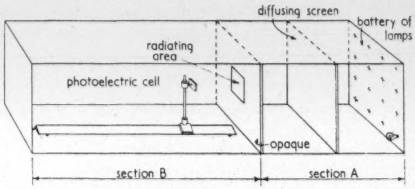


Fig. 1. Apparatus for determination of the configuration factor

paper and covered with a diffusing layer of material is pasted on to the transparent plastic screen which is uniformly illuminated from behind. The inside of section *A* is painted white to give a high reflectivity. The radiating area is uniformly illuminated and the light radiation from it obeys Lambert's cosine law. The inside of section *B* is painted black to prevent reflected light from falling on the calibrated barrier layer photocell. The cell, with a sensitive area of about 1 cm^2, is used in conjunction with a galvanometer to measure the light intensity. If the intensity of the light source can be maintained constant, for instance by incorporating a constant voltage transformer, a calibrated scale can be fitted to the galvanometer to indicate the configuration factor directly. Thus when the photocell is placed close to the radiating area the configuration factor will be unity, and when placed in the position of the receiving element, the configuration factor can be read directly from the scale. Should there be reflecting surfaces in the heat problem between the radiating surface and the point at which the measurement is to be made, these, with suitable reflectivity coefficients, would have to be included in the illuminated model.

Having set up the model of the radiating area, the configuration factor can be quickly found at any position with

respect to the radiator by merely placing the photocell in that position and reading off the configuration factor on the scale. An idea of the laborious computation which is required to solve even the simple case of a receiving surface parallel to and on the axis of a rectangle can be seen from the expression for the configuration factor

$$\phi = \frac{2}{\pi}\left[\frac{x}{\sqrt{(x^2+y^2)}}\tan^{-1}\frac{z}{\sqrt{(x^2+y^2)}} + \frac{z}{\sqrt{(z^2+y^2)}}\tan^{-1}\frac{x}{\sqrt{(z^2+y^2)}}\right]$$

in Fig. 2 (see inset), which shows the accuracy of the method outlined in this paper. The values of configuration factor near to the radiator, given by the photometric method, are lower than the calculated values. This is not illustrated well by Fig. 2, since for small values of distance the curve is ill-

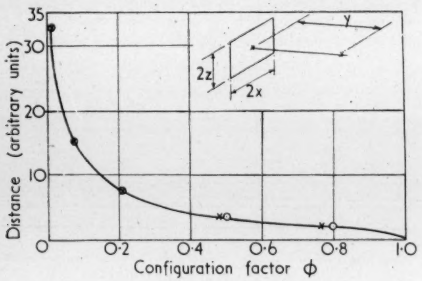


Fig. 2. Configuration factor at various distances along axis of rectangular source having dimensions $6 \times 9{\cdot}14$ (arbitrary units)

× = experimental points; ○ = calculated points.

conditioned with respect to the ϕ axis. The reason for the low values is that the cell receiving illumination at a large angle of incidence gives a response lower than the correct value, which should be proportional to the cosine of the angle of incidence. The deviation from the theoretical response is commonly known as the "cosine error" of the cell, and a simple method of correcting this error has recently been described by Pleijel and Longmore.[6] With this modification to the photocell the values of configuration factor near to the radiator could be found with the same accuracy as those some distance from the radiator.

APPLICATION OF THE PHOTOMETRIC METHOD TO PROBLEMS OF TOTAL HEAT TRANSFER BY RADIATION

Problems in furnace design and panel heating, as well as many others, often require a knowledge of the total heat transfer between two finite bodies, and not the amount of heat falling on a small element from a finite source. In this case the configuration factor of the whole receiving element with respect to the radiator is required. This is known as the integrated or mean configuration factor, $\bar{\phi}$, and can be obtained from the expression for ϕ in the Appendix by integration over the area A_2. If we consider it as the mean configuration factor, that is to say the mean of the values of configuration factor for elements distributed all over the area A_2, then it can be expressed as

$$\bar{\phi} = \frac{1}{A_2\pi}\int_{A_1}\int_{A_2}\frac{\cos\theta_1\cos\theta_2}{x^2}dA_1dA_2$$

A close approximation can be made for the mean configuration factor ($\bar{\phi}$) of a large receiving surface by taking the configuration factor, ϕ, at a few points on it in suitable positions with the photocell and taking the average of these values.

Many problems involve the transfer of heat from several surfaces. It is useful to remember in these cases that configuration factors can be added and subtracted and the configuration factor can thus be found separately for each surface, the total configuration factor being given by the sum of the values.

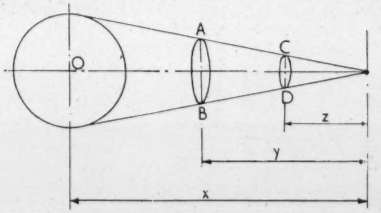


Fig. 3. Configuration factor of sphere and equivalent radiating surfaces

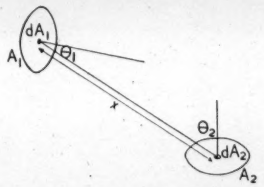


Fig. 4. The transmission of radiation between two surfaces A_1 and A_2

A solution of particular cases in which the radiator is three-dimensional can be obtained using the photometric method, since it is the contour of a surface and not the form which determines the value of ϕ. Thus the surface of a radiator may be replaced by an equivalent surface, provided the angular cover at the receiver is the same. This can easily be seen by considering the configuration factor of a sphere at a distance x (Fig. 3).

The value of ϕ at a point distant x from the centre of the sphere centred on O, would be identical with the value for the disk AB at a distance y, or the disk CD at a distance z.

CONCLUSIONS

The photometric method of determining configuration factors described in this paper is particularly useful if the configuration factor at many positions with respect to a radiator is to be found. It is also possible by this method to find at what position with respect to a radiator there is a

given configuration factor. The authors know of no other method of solving this problem for any but the simplest cases.

ACKNOWLEDGEMENTS

The work described in this paper forms part of the programme of the Joint Fire Research Organization of the Department of Scientific and Industrial Research and Fire Offices' Committee; the paper is published by permission of the Director of Fire Research.

REFERENCES

(1) BARKER, A. H., and KINOSHITA, M. *University of London, University College Department of Heating and Ventilating Engineering Bulletin*, No. 1 (London: Unwin Bros., Ltd., 1923).
(2) MCGUIRE, J. H. *Heat transfer by radiation.* Department of Scientific and Industrial Research and Fire Offices' Committee Fire Research Special Report No. 2. To be published. (H.M. Stationery Office.)
(3) HERMAN, R. A. *Treatise on Geometrical Optics.* (Cambridge: Cambridge University Press, 1900).
(4) ECKERT, E. *Z. Verein. Dtsch. Ingen.* [*VDI*], **79**, p. 1495 (1935).
(5) HOTTEL, H. C. *Mech. Engng*, **52**, p. 699 (1930).
(6) PLEIJEL, G., and LONGMORE, J. *J. Sci. Instrum.*, **29**, p. 137 (1952).

APPENDIX

The intensity of radiation at any point with respect to a radiator can be expressed as

$$I = \phi\epsilon\sigma T^4$$

where ϕ is the configuration factor.

The fundamental expression for ϕ can be deduced from Fig. 4.

The energy falling on dA_2 in unit time from the element dA_1 of the radiator is

$$dA_1 \frac{\epsilon\sigma T^4}{\pi} \cos\theta_1 \times \frac{dA_2 \cos\theta_2}{x^2}$$

Considering this as the intensity of radiation at dA_2 from the whole radiator

$$I = \frac{\epsilon\sigma T^4}{\pi} \int_{A_1} \frac{\cos\theta_1 \cos\theta_2}{x^2} dA_1$$

Therefore

$$\phi = \frac{1}{\pi} \int_{A_1} \frac{\cos\theta_1 \cos\theta_2}{x^2} dA_1$$

The thermionic emission from thin films of barium and strontium oxide

By J. WOODS, Ph.D., A.Inst.P., and D. A. WRIGHT, M.Sc., F.Inst.P., M.-O. Valve Co. Ltd. at the G.E.C. Research Laboratories, Wembley, Middlesex

[*Paper first received* 5 *June, and in final form* 28 *July*, 1953]

The thermionic emission from thin evaporated films of barium and strontium oxide has been studied. Thin barium oxide films have a good emission following evaporation. The emission improves a little during heat treatment, and is then similar to that from a sprayed cathode coating of barium oxide. There is a maximum emission at a thickness of 10^{-5} cm. Strontium oxide behaves in a similar way, with a lower level of emission. If barium oxide is evaporated on strontium oxide, the best emission obtainable is as for barium oxide alone. If, however, strontium oxide is evaporated on barium oxide, heat treatment results in an emission higher than that from either barium or strontium oxide alone, and similar to that from a sprayed cathode coating of mixed barium and strontium oxide. The emission is best when the strontium oxide thickness is 10^{-5} cm, whether the barium oxide thickness is of this order or greater.

In previous publications[1,2] the authors have described some of the thermionic and secondary emitting properties of evaporated films of barium oxide, with the chief emphasis on secondary electron emission under electron bombardment. In this paper further studies of the thermionic emission of barium oxide are presented. The emission of films of strontium oxide, and of strontium oxide on barium oxide as well as barium oxide on strontium oxide has also been studied in an attempt to add to the knowledge of the mechanism of thermionic emission from mixed-oxide cathodes.

EXPERIMENTAL ARRANGEMENT

The thermionic emission of the evaporated layers was compared with the emission of sprayed coatings of barium oxide, strontium oxide and mixed barium and strontium oxide. All measurements were made using the same type of diode, in which an 8 W box cathode of either pure or O-nickel was mounted on a cylindrical framework which could be rotated. O-nickel is nickel to the following specification: nickel + cobalt, 99·5% minimum; magnesium, 0·07–0·15%; cobalt, $< 0·5\%$; manganese, $< 0·15\%$; iron, $< 0·2\%$; silicon, $< 0·1\%$; copper, $< 0·1\%$; sulphur, $< 0·005\%$. The alkaline earth oxide was evaporated from a platinum helix previously coated with the appropriate carbonate, which was converted to oxide by heating at 1000° C during pumping. When the oxide was to be evaporated the cathode faced the helix; the cathode was then 180° from the position where it faced the plane rectangular sheet of nickel which formed the anode.

With the valves in which the emission of sprayed coatings was to be examined, the cathodes were sprayed with the appropriate carbonates before assembly. Coating thicknesses from 0·04 to 0·1 mm were used. Evaporating coils were not provided in these valves.

The valves were pumped on a mercury diffusion pump

with a liquid air trap. After baking at 400° C for ½ h, all metal parts were outgassed by eddy-current heating. The carbonates were converted to oxides and all metal parts outgassed again. The valves were then sealed from the pump and gettered. During the conversion of the carbonate coating on the platinum helix, the cathode faced the anode, thus avoiding contamination with the material evaporating from the helix at this stage.

Valves with sprayed coatings were aged by maintaining the cathodes at 850° C for 20 min while a space-charge limited current of 60 mA/cm² was drawn. With a diode with a platinum helix a considerable amount of oxide was distilled on to the surroundings before the cathode was rotated to face the evaporating coil and films were formed on the cathode. The evaporation of barium oxide was carried out at 1200° C, and of strontium oxide at 1400° C. At these temperatures the evaporation of the respective oxides proceeded at roughly equal rates. Measurements of both d.c. and pulsed emission were made. The pulsed emission was determined using pulses of 2 μsec duration with a repetition rate of 50 p.p.s. The d.c. emission was measured at temperatures up to 500° C where it was stable. At temperatures near 600° C and higher the d.c. emission could not be determined owing to rapid emission changes with time.

Unless otherwise stated all temperatures quoted are brightness temperatures. Cathode temperatures in the range 700 to 1000° C were measured with a disappearing filament pyrometer, while temperatures between 500 and 700° C were determined by extrapolation of the plot of cathode brightness temperature against heater wattage. With evaporated films the pyrometer was sighted on the nickel cathode before the evaporation of the oxide, while with the sprayed coatings the pyrometer was focused on the nickel at the side of the cathode immediately next to the coating.

BARIUM OXIDE

The thermionic emission of evaporated barium oxide has been measured for a variety of thicknesses between 10^{-6} and 10^{-4} cm. The thickness of a layer was estimated from the interference colours formed by reflexion of white light. The results are recorded in the table and were obtained with films deposited on O-nickel. In general the behaviour of films deposited on pure nickel was very similar although the emission level was about one-half of that on O-nickel at all thicknesses.

The pulsed emission of a film as evaporated was usually about half the emission obtainable when the film was fully activated. Activation was achieved by taking pulsed emission at 800° C for 15 min. The emission also varied with the thickness of a layer; it was lowest for thicknesses of 10^{-6} cm and increased to an optimum value for thicknesses near 10^{-5} cm. As the deposition of barium oxide was continued the maximum emission obtainable from a film gradually fell.

With the platinum helix held at 1200° C, films 10^{-6} cm thick were formed in two or three minutes. The d.c. emission of such a film was measured with the cathode at 500° C. After standing for 30 min at this temperature the emission usually improved by about 50%. At the higher temperatures where the pulsed emission was determined thermionic activation also took place. The process was slow between 700 and 800° C, but above 800° C a state of maximum emission was quickly reached. If these temperatures were maintained for too long, however, the emission began to fall. This was due to the evaporation of material from the cathode.

With thicker films the behaviour was essentially similar to that just described. At 10^{-5} cm the d.c. emission of a film as deposited was some three times higher than the d.c. emission of a thinner film of 10^{-6} cm immediately after its formation. Thermionic activation was effected by taking pulsed emission at 820° C, and although evaporation of the oxide from the cathode took place at this temperature the film was thick enough for the emission to be maintained over periods of hours. At greater thicknesses of up to 10^{-4} cm, both the d.c. emission of a freshly deposited layer and its activated (pulsed) emission were slightly less than the corresponding emission for a film of the optimum thickness.

After activation the saturated pulsed emission i_s was determined over the temperature range 700° to 850° C, and Richardson plots were made by plotting $\log i_s/T^2$ against $1/T$. The brightness temperatures were converted to true temperatures by assuming a value of 0·35 for the spectral emissivity of nickel at a wavelength of 0·665 μ. In the temperature range studied the Richardson plots yielded straight lines. The work function determined from the slope of these lines was near 1·2 eV for all films, no matter what the thickness. For

Thermionic emission from films and coatings of barium and strontium oxide

Cathode	*Thickness* cm	*Emission as deposited*		*Activated emission*	
		d.c. at 550° C	*pulsed at 755° C*	*d.c. at 550° C*	*pulsed at 755° C*
BaO	10^{-6}	0·67 mA/cm²	0·2 A/cm²	1·2 mA/cm²	0·4 A/cm²
	10^{-5}	2·0	0·6	3·0	1·2
	10^{-4}	1·5	0·5	2·0	0·9
	10^{-2} (spray)	—	—	0·4	0·6
SrO	10^{-6}	0·002 mA/cm²	—	—	0·1 at 860° C
	10^{-5}	0·4	0·6 at 860° C	—	0·1 at 755° C 1·4 at 860° C
	10^{-4}	0·3	0·4 at 860° C	—	0·9 at 860° C
	10^{-2} (spray)	—	—	0·002	0·14 at 860° C
SrO on BaO	both 10^{-5} cm thick	0·008	0·3	9·0	3·0
(BaSr)O	10^{-2} cm (spray)	—	—	4·0	3·5
BaO on SrO	both 10^{-5} cm thick	—	0·1	—	0·6

comparison purposes the emission of well activated, sprayed coatings of barium oxide on O-nickel was determined. The emission at 755° C was 0·6 A/cm² and the work function deduced from a plot of $\log i_s/T^2$ against $1/T$ was 1·4 eV. The best emission from the barium oxide films was 1·2 A/cm² at 755° C, as shown in the table.

STRONTIUM OXIDE

Strontium oxide was evaporated on to O-nickel cathodes by heating the oxide-coated platinum helix at 1400° C. At this temperature, films 10^{-5} cm thick were deposited in about an hour. The thermionic properties of thin films of strontium oxide were very similar to those of evaporated barium oxide, although the emission levels were different.

As with barium oxide, thermionic activation took place when pulsed emission was drawn at 850° C. At higher temperatures so much material was lost from the cathode by evaporation that the layer was rapidly de-activated. The emission of the strontium oxide films varied with thickness in exactly the same way as the emission of barium oxide films. The maximum emission was obtained at a thickness of 10^{-5} cm, and was considerably greater than that from a sprayed coating (see the table). The work function of the evaporated films of strontium oxide calculated from Richardson lines in the temperature range 750° to 900° C was near 2·0 eV for all thicknesses. The corresponding value for a sprayed coating was 3 eV in the present experiments, although values near 2 eV have been obtained in earlier work.

MIXED LAYERS OF BARIUM AND STRONTIUM OXIDES

Measurements have been made on composite layers prepared by evaporating strontium oxide on top of previously evaporated films of barium oxide and *vice versa*. The barium oxide film was put down first and activated. After deposition of some strontium oxide the d.c. emission at low temperatures was very small; at temperatures above 750° C, however, thermionic activation took place and became very rapid between 800 and 850° C. The activated emission current was greater than that from the original barium oxide film.

In these experiments the initial barium oxide films were of two thicknesses, 10^{-5} and 10^{-4} cm. As strontium oxide was deposited on top of these films the emission obtainable increased to an optimum value when the thickness of the strontium oxide film reached 10^{-5} cm. The behaviour was the same for both thicknesses of the underlying barium oxide. As more strontium oxide was deposited the available emission gradually fell. The best emission measured was 3·0 A/cm² at 755° C for a film of strontium oxide 10^{-5} cm thick on top of a film of barium oxide of the same thickness. This compares favourably with the value of 3·5 A/cm² for a well activated sprayed coating of mixed barium and strontium oxide at the same temperature. This result indicates clearly that the emission from composite evaporated layers is a function of the quantity of strontium oxide on the surface and not of the relative proportions of the two oxides.

Richardson plots of $\log i_s/T^2$ against $1/T$ in the range 650° to 800° C for layers of strontium oxide on barium oxide of different composition all yielded straight lines. The work function of the most emissive layers obtained from the slopes of these lines was 1·2 eV.

Equally good emissive layers were not obtained when strontium oxide was put down first and barium oxide was evaporated on top of it. Evaporation of barium oxide on the original well activated strontium oxide film caused a considerable drop in the low temperature emission. Once again activation was accomplished by drawing pulsed emission at 800 to 850° C. However, the best emission obtainable was not comparable with that of strontium oxide on barium oxide. In fact as the deposition of barium oxide proceeded the activated emission merely increased from that of pure strontium oxide to that of pure barium oxide. If after prolonged evaporation of barium oxide, strontium oxide was now deposited, the activated emission increased to a maximum, and in all respects the behaviour was identical with that described in the preceding paragraph.

When barium and strontium oxides were evaporated simultaneously at equal rates, the activated emission was not as high as that for a layer of strontium oxide 10^{-5} cm thick deposited on barium oxide, but was nevertheless slightly higher than the emission of barium oxide.

DISCUSSION

The variation with thickness of the thermionic emission of films of barium oxide evaporated from platinum was discussed in the earlier papers[1,2] in terms of lattice vacancies and crystallization effects. No further evidence has been obtained to show whether the qualitative explanation given there is correct.

The results with mixed layers of barium oxide and strontium oxide are instructive in connexion with the mechanism of the oxide cathode. Similar results were obtained by Huber[3] working with sprayed coatings, in that he found that a thin layer of strontium oxide sprayed on a thicker layer of barium oxide gave an emission greater than that of the barium oxide alone. It has been shown[4] that the oxide cathode with barium and strontium oxides in roughly equal proportions rapidly develops a surface layer consisting mainly of strontium oxide. That this surface of strontium oxide was instrumental in raising the emission of the double oxide above that of barium oxide was proposed in 1944 by Herrmann and Wagener,[5] and a similar suggestion was made independently by one of the writers later.[6] The present results strongly support the view that to obtain the high emission characteristic of the oxide cathode, it is necessary to have a surface layer of strontium oxide. This is not effective, however, unless the cathode has been heated above 750° C; this process presumably permits barium originating in the barium oxide to diffuse outwards into the strontium oxide, and perhaps on to its outer surface.

REFERENCES

(1) Woods, J., and Wright, D. A. *Brit. J. Appl. Phys.*, **3**, p. 323 (1952).

(2) Woods, J., and Wright, D. A. *Brit. J. Appl. Phys.*, **4**, p. 56 (1953).

(3) Quoted by Herrmann, G., and Wagener, S. *The Oxide Cathode*, Vol. II, p. 230 (London: Chapman and Hall, Ltd., 1951).

(4) Gärtner, H. *Phil. Mag.*, **19**, p. 82 (1935).
Darbyshire, I. A. *Proc. Phys. Soc.* [*London*], **50**, p. 635 (1938).
Huber, H., and Wagener, S. *Z. Tech. Phys.*, **23**, p. 1 (1942).
Eisenstein, A. *J. Appl. Phys.*, **17**, pp. 434, 654 (1946).

(5) Herrmann, G., and Wagener, S. *Die oxyd cathode*, Vol. II, p. 188 (Leipzig: Barth, 1944).

(6) Wright, D. A. *Proc. Phys. Soc.* [*London*], **60**, p. 13 (1948).

New books

Introduction to solid state physics. By PROF. CHARLES KITTEL. (London: Chapman and Hall Ltd.; New York: John Wiley and Sons, Inc.) Pp. xii + 396. Price 56*s*.

The physics of the solid state is now so large a subject that to cover it all in one textbook makes drastic selection necessary. Professor Kittel, by his own experimental work, especially on nuclear resonance, has himself made very important contributions to the subject, and, as one would expect from him, the hallmark of the book is modernity. It is definitely a textbook; it is meant for university students, and examples and problems are included at the end of each chapter; but the subjects chosen, though always treated in an elementary way, are nearly all those in which important work has been done in the last five years. Examples are: the thermal conductivity of solids, with a discussion of the size effect; ferroelectrics; nuclear and electronic spin resonance absorption; antiferromagnetism; transistors and the dislocation theory of crystal growth. Obviously none of these subjects can be treated very thoroughly, but in each case there is an elementary introduction, some account of the main results and references to some of the more important summarizing papers. The book should be very useful in British universities, especially those that do not specialize in the physics of the solid state, to show both students and staff what is really going on in this subject, and to allow them to judge what could be included in a comprehensive physics course.

N. F. MOTT

Mechanical properties of wood and paper. By W. W. BARKAS, R. F. S. HEARMON, A.Inst.P., and H. F. RANCE. (Amsterdam: North-Holland Publishing Co.) Pp. xii + 298. Price 50*s*.

This book forms one of a series of monographs on the rheological behaviour of natural and synthetic products. It gives a wide review of published work on the mechanical properties of wood and paper.

The first part, by Barkas and Hearmon, occupies just over a third of the volume and deals with wood. Topics introduced include the structure of wood, its elastic and plastic properties, hygroscopicity, swelling stresses, plastic swelling and sorption hysteresis. Much of this information is now easily accessible, but it is useful to have this clear, concise account of it.

The remainder of the book, by Rance, is concerned with paper, its manufacture, testing, hygrosensitivity and the effect of its structure on its physical properties. The study made covers a great deal of ground and is ably done, but one is left in some doubt at the end of it all whether the labour has been worth while. In the author's words, the treatment is "discursive and largely speculative, on account of the relative paucity of experimental records in the field of paper research. Published opinions are more varied and numerous than experimental studies." As a result, one has only to read a score or so of pages to get hopelessly bogged in a morass of qualitative theories and pseudo-scientific jargon with little hope of rescue by a neat piece of quantitative experimental work. A less detailed and more discriminating review might enable us to see whether or not there is a Wicket Gate beyond the Slough of Despond.

It may be, however, that this second part of the volume will induce some rheologists to forget about micelles, hydrogen bonds, primary and secondary creep, micro-crêping, fibrillation, flocculation and the other weapons in the theorist's armoury, and to concentrate for a year or two on exact quantitative measurement. In that event, Dr. Rance's work will have been well worth while. Paper rheology offers a promising and interesting field of research but at present we suffer from over-abundance of speculation and a dearth of plain physics.

The volume is well produced and printed. Mistakes and misprints are few and unimportant. V. G. W. HARRISON

Grössengleichungen, einheiten und dimensionen. By JULIUS WALLOT. (Leipzig: Johann Ambrosius Barth.) Pp. viii + 215. Price D.M.16.35.

The three main sections of this book deal with so-called "magnitude equations," units and dimensions, as indicated in the title.

According to the author the physical laws are expressed in the form of equations between "magnitudes" as distinct from "Zahlenwertgleichungen," i.e. equations between the numerical values of these magnitudes with respect to a definite system of units. But, as the basic algebraic operations are defined for numbers only, it does not seem possible to connect any definite meaning with, say, the product of "mass" and "acceleration," unless the operation "multiplication" is clearly defined for these physical magnitudes. The mathematician is, of course, at liberty to define general algebraic operations for abstract mathematical quantities which are not numbers (e.g. matrices or operators); but relationships of this kind will represent physical laws only if they are verifiable, i.e. if numerical relationships can be uniquely deduced from them which can be compared with relations between numerical results of measurements. This whole section is therefore rather unsatisfactory and many points are open to objection.

The second section deals with the thorny subject of the various systems of units in physics. The definition of the primary and secondary units in these systems is given with great care and in considerable detail and the merits and shortcomings of the systems are discussed. The many tables presenting the connexions between the most important units in the various systems will be found very valuable in view of the fact that no two authors nowadays seem to be using the same system of units. The use of the term "rational" for one group of systems of units is regrettable, as there is no reason why any arbitrary convention concerning numerical factors in certain formulae should be more "rational" than any other.

The last section gives a clear presentation of the meaning of "dimensions" of physical quantities and an introduction into the method of dimensional analysis and the so-called "model theory." A selection of examples from different fields of physics is given which demonstrate the usefulness of these methods under certain conditions, but also show that they can be misleading unless great care is taken in the selection of those quantities which are relevant to the problem in question.

The book is written in a lively and unconventional style which, however, carries with it a certain amount of superficiality and vagueness. Physicists engaged in the application of the subject will certainly find it very useful. R. FÜRTH

Microwave spectroscopy. By W. GORDY, W. V. SMITH and R. F. TRAMBARULO. (London: Chapman and Hall, Ltd.) Pp. xii + 446. Price 64*s*.

The great progress in microwave technique made during the war has led to far-reaching results in the study of atomic nuclei, atoms, molecules and solids. To mention only a few: the shift of the $2^2S_{\frac{1}{2}}$ level of atomic hydrogen discovered with the help of microwaves (the so-called Lamb shift) is of extreme importance in quantum electrodynamics; nuclear spins can be found; the angular momentum of electrons in paramagnetic and ferromagnetic solids can be obtained (g factor); details of molecular structure can be measured, and numerous applications to chemical and to isotopic analysis can be made.

It is clear that a book of moderate size cannot possibly introduce in detail all these subjects. Since the authors' aim was to survey the whole subject in a single book they had to restrict themselves severely. This might not appeal to the specialist but it greatly helps the non-specialist to obtain information on the possibilities of this new experimental technique.

The book is divided into nine sections dealing with experimental methods and the various applications, together with a survey of the relevant properties of nuclei, atoms and molecules.

This appears to be the first book on the subject, and the extensive bibliography must be welcomed by all those who wish to make use of the technique. The very extensive appendix of atomic and molecular constants increases the usefulness of the book.

H. FRÖHLICH

Measurement of humidity. (London: H.M. Stationery Office.) Pp. iv + 18. Price 1*s*.

This pamphlet is No. 4 of the series "Notes on applied science" issued by the National Physical Laboratory with the object of providing for industrialists and technicians information on various scientific and technical subjects which is not readily available elsewhere." It is on the most elementary level, and concentrates on the aspirated psychrometer and dew point hygrometer. There is an unfortunate lack of clarity in the treatment of measurements below 0° C: we read on p. 3 "Below 0° C the dew point is termed the 'hoar frost point.' "

It is surprising that "industrialists and technicians" should today be considered in need of instruction at this level; may it not be that an author who can write "A useful substitute [for wet bulb sleeving] is cotton corset lace with the central strands removed" is a little out of touch with the times?

G. D. ROBINSON

Technical aspects of sound. Edited by E. G. RICHARDSON, B.A., Ph.D., D.Sc. (Amsterdam: Elsevier Publishing Co.; London: Cleaver-Hume Press Ltd.) Pp. xviii + 544. Price 70*s*.

This is the first of two volumes intended to provide, for the first time, a "handbook" surveying all the technical aspects and applications of sound. This volume deals with the audible range of airborne sound, including divisions on fundamentals, architectural and physiological acoustics, noise, recording and reproduction, and musical instruments. The second volume will be concerned with ultrasonics and underwater sound.

Each of the nineteen chapters presents a compact account of the theory and practice of its subject and duplication of the same material by two or more of the seventeen authors has been almost entirely avoided. Many of the contributions are excellent, special praise being due to the general introduction, the chapters on absorbing materials, speech and hearing, and the fascinating section on musical instruments. That on sound recording leaves most to be desired, particularly the part dealing with magnetic recording, which is spoilt by too great a preoccupation with the details of certain proprietary equipment. It is felt also that the omission from the chapter on architectural acoustics of any mention of the special problems of broadcasting and recording studios is a serious defect, as it is these very problems which have stimulated much of the acoustical research of the last thirty years.

With these reservations the book provides extremely valuable general reading for the advanced student or research worker.

C. L. S. GILFORD

Proceedings of the Western Computer Conference, held by the Joint IRE–AIEE–ACM Computer Conference Committee. (New York: The Institute of Radio Engineers Inc.) Pp. 231. Price $3.50.

Reports of conferences organized by specialist groups now form an important part of periodical or semi-periodical scientific literature. Those interested in computing machinery will be familiar with reports of conferences organized jointly by the Institute of Radio Engineers, the American Institute of Electrical Engineers, and the Association for Computing Machinery. Hitherto these reports have referred to conferences held on the east coast of the United States; the present volume is a report of the first of a series of conferences, necessarily of a less ambitious nature, that it is proposed to hold under the same auspices on the west coast.

The conference reported on was held in Los Angeles in February 1953. The various sessions were concerned with business applications of digital computers, the evaluation of analogue and digital computers, applications of analogue computers to special problems in aircraft design, components for digital computers, flight simulators, and analogue-to-digital converters. There was one paper on the use of an electronic differential analyser to solve partial differential equations.

M. V. WILKES

Fluorescence of solutions. By E. J. BOWEN and F. WOKES. (London: Longmans, Green and Co. Ltd.) Pp. 91. Price 25*s*.

This book is intended to assist those interested in fluorescence as a tool for analysis; for example, in assays of the B vitamins. Its short length makes the various treatments of fundamental phenomena, fluorescence detection and details of analysis methods somewhat sketchy. There are five chapters on fundamentals such as conditions for fluorescence, absorption and emission spectra, mechanisms of quenching, etc., with three other chapters covering the apparatus used in detection, e.g. light sources, filters, solution cuvette design, photocells and commercial fluorimeters. Mere mention is made of the modern photosurfaces available such as those of the antimony or bismuth types with characteristics superior to others described in more detail. The discussion of apparatus is somewhat uncritical, e.g. no mention is made of the serious temperature drift of barrier layer photocells or of the spatial instability of mercury arc sources. Two final chapters are concerned with actual methods of analysis, attention being confined to assay of aneurin and riboflavin which are of considerable commercial importance. For its size the book is rather expensive and it will not satisfy the demand for a sound and comprehensive treatment of the whole field of the application of fluorescence to chemical analysis.

G. F. J. GARLICK

Crystal growth and dislocations. By A. R. VERMA, M.Sc., Ph.D. (London: Butterworths Scientific Publications.) Pp. xii + 182. Price 30*s*.

This book is a survey of recent work on crystal growth, with special reference to spiral dislocations. It includes chapters on the theory of crystal growth, types of imperfection in crystals, experimental methods, and practical results. In fact, however, the emphasis is almost entirely on spiral dislocations, and the book contains a beautiful collection of over forty photographs of them.

Apart from these photographs, the value of the book is somewhat doubtful; it seems to the reviewer that the subject is too new for it to deserve treatment in a textbook. (On p. 54, the author says that the first relevant experimental observations were made in 1950.) Moreover, even if such treatment were deserved one would welcome it from a person with more background experience of the subject of crystal growth; it would appear that the author was writing this book and his Ph.D. thesis at the same time.

In fact, the book has many of the deficiencies of a Ph.D. thesis. In places (e.g. parts of Chapter 8) it degenerates into a mere catalogue; some subjects (e.g. the Burgers vector, p. 24) are mentioned before they are formally defined; and the treatments of experimental methods (e.g. interferometry) are rather perfunctory. The treatment of phase-contrast microscopy is not at all clear, and the diagram illustrating it (Fig. 11) seems to be incorrect.

But the photographs are beautiful. H. LIPSON

An advanced treatise on physical chemistry. Vol. 4. By J. R. PARTINGTON. (London: Longmans, Green and Co. Ltd.) Pp. xv + 688. Price 80*s*.

The fourth volume of Professor Partington's *Advanced treatise on physical chemistry* is more unified in its range than any of the previous volumes. It deals with various aspects of the subject which the author calls "physico-chemical optics." The matter, selected from a wide field, includes many items of special interest to chemists.

In a volume of approximately 700 pages, the division of subject matter is approximately as follows: refraction (1/7), polarization (2/7), optical activity (1/7), electro-magnetic theory of light (2/7). Short monographs on magnetic rotation (41 pages), piezo- and pyro-electricity (7 pages), together with a mathematical appendix makes up the remainder. The properties of lenses and diffraction gratings as well as measurement of spectra are omitted on the grounds that they are "of more physical interest and are dealt with in many available treatises". However, an elementary account of molecular spectra is promised in the next volume. A similar account of microscopy and related topics—especially on the experimental side—would also be welcome.

The author gives a systematic and detailed survey of the field with that lucidity and thoroughness which have been characteristic of the previous volumes. Over 10 000 references are cited; many date back to the last century, but several play an important part in giving an over-all picture of the matter covered. Another feature of this volume is the large number of clear, well-lettered line diagrams, essential to any treatise on optics. The scarcity of errata in such a large volume is also noteworthy.

The usual comprehensive scientific text, compiled in collaboration with a number of experts, often lacks co-ordination and uniformity of treatment. Partington's treatise, being the work of one man, does not suffer from these defects.

JOSEPH REILLY

Styrene. By R. H. BOUNDY and R. F. BOYER. (London: Chapman and Hall Ltd.) Pp. xxii + 1304. Price 160*s*.

This book is a monograph on styrene and describes its manufacture, its polymers and copolymers and a large number of its chemical derivatives. It comprises twenty-two sections which were compiled by twenty-five authors. The two editors have world-wide reputations for their work on styrene and its polymers and they, and most of the authors, are members of the staff of the Dow Chemical Company of the U.S.A. This company commenced work on styrene about 1930 and established a styrene production which was increased enormously from 1942 onwards because of the war-time requirements for styrene-butadiene synthetic rubber. The interest of this book lies not only in its technical content, but also in the picture it gives of the approach of a larger American chemical company when considering the manufacture and applications of a material. The scope of this approach is extremely wide and illustrative of the large resources necessary for the successful exploitation of materials in the present industrial age.

It is, of course, practically impossible for one individual to give an adequate review of a work covering such a wide field. A list of the sections would be almost a review in itself. However, it can be assumed that readers of the present journal have at least a bias towards the physics of the materials discussed in the monograph. Certain sections should be of particular interest for them. Thus, the large part of the book which is quite correctly given to the discussion of polystyrene, contains much valuable information. Perhaps the outstanding section is that dealing with the molecular weight. This is a fine survey of the methods which can be applied to thermoplastics in general with, of course, special reference to polystyrene. A section on the general physical properties of polystyrene is informative, but rather disappointing in its meagre treatment of both impact strength and creep. The optical and electrical properties of polystyrene are treated adequately and include an interesting discussion on the extremely low dielectric absorption of this material. The section covering these properties is followed by one on the rheology of polystyrene. This is comparable, in its quality and applicability to thermoplastics in general, with that cited above in connexion with the molecular weight. Finally, those interested in the manipulation of thermoplastics will find much of interest in the descriptions of the processes used for the fabrication of articles from polystyrene. The above attempt to interpret the interests of the readers of this journal must not be taken as a denigration of the many other sections. It is felt that other reviewers will consider the treatment in the monograph of the kinetics of styrene polymerization and copolymerization; the chemical modifications; and the various styrene resins, to say nothing of the extremely large bibliography of the patents which have been written on co-polymerization.

This book can certainly be recommended as a reference book for those interested in styrene, its polymers and derivatives. Workers in the field of thermoplastics will also find it useful as a general guide to a number of the physical considerations of the behaviour of their materials. Its price puts it beyond the reach of most individuals, but the book should prove an asset to either the academic or the industrial science library. The standard of printing and production is of the usual American excellence. Such typographic errors as have been noticed were obvious and would presumably be corrected in a later edition.

H. A. NANCARROW

Notes and comments

Elections to The Institute of Physics

The following elections have been made by the Board of The Institute of Physics:

Fellows: K. F. Bowden, R. N. Bracewell, P. A. Feltham, R. F. S. Hearmon, F. J. Kerr, A. Quinton, B. Sugarman, N. Veall.

Associates: A. A. Abou-el-azm, J. Bagot, J. E. Barrow, J. Bennett, S. J. Carlisle, A. J. L. Collinson, J. W. C. Gates, E. Gillam, D. L. Goldby, I. H. Hardwich, A. G. Hester, D. G. Jaquess, W. G. P. Lamb, C. A. Mann, I. A. Marshall, K. G. Matthews, H. J. Mead, R. H. A. Miles, A. E. Moore, J. J. O'Dwyer, D. H. Pringle, J. Stockill, J. W. Sturgess, M. Thomas, G. A. J. Vandermeerssche.

Eighteen Graduates, fifteen Students and eleven Subscribers were also elected.

The Physical Society's annual exhibition of instruments and apparatus

The Physical Society's 38th annual exhibition of scientific instruments and apparatus will be held at South Kensington, London, from 8 to 13 April next (excluding Sunday).

Tickets for admission may be obtained from The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. Members of The Institute of Physics may obtain tickets from the Institute's office.

As usual, a handbook, price 7*s.* 3*d.* (including postage), containing brief descriptions of the exhibits will be issued by the Society about a month before the exhibition. The entries submitted to the craftsmanship and draughtsmanship competition for apprentices and learners will be displayed at the exhibition.

Improvement of electronic components

The Institute of Radio Engineers has organized a Component Parts Professional Group for the purpose of promoting continued improvement of electronic components and the exchange of information. This group's field of interest includes the characteristics, limitations, applications, development, performance and reliability of component parts. With the co-operation of the American Institute of Electrical Engineers, other professional groups of the IRE, the Radio-Electronics Television Manufacturers' Association, the West Coast Electronics Manufacturers' Association, the U.S. Department of Defense and the National Bureau of Standards, the group will sponsor the Electronic Components Symposium to be held at the Department of Interior in Washington, D.C., U.S.A., on 4–6 May, 1954.

Other activities planned by the group are the publication of a transaction in the *Proceedings of the Institute of Radio Engineers* in 1954, and a scholarship plan to induce suitable graduates to direct their activities towards the development of electronic components.

Application forms and detailed information about this group may be obtained from Dr. Fred Haynes, Supervisor, Electronics Product Section, The Glenn L. Martin Co., Baltimore, Maryland, or application may be made direct to Miss Emily Sirjane, Institute of Radio Engineers, 1 East 79th Street, New York 21, N.Y., U.S.A.

International conference on semiconductors

The Netherlands Physical Society, with the support of the International Union of Pure and Applied Physics, and UNESCO will organize an International conference on semiconductors, to be held at Amsterdam from 29 June to 3 July, 1954.

Subjects to be discussed will include bulk recombination; surface conductivity; surface trapping; surface recombination; intermetallic compounds, the band picture in polar and non-polar semiconductors; photoconductivity in semiconductors like PbS, PbTe, PbSe, ZnS, CdS; the application of general physical and chemical laws to the preparation of semiconductors with specific properties.

In order to avoid the overlapping of the topics dealt with at this conference with those to be dealt with at the Conference on defects in crystalline solids to be held at Bristol from 12 to 17 July, 1954, the Organizers of both conferences have chosen the topics so that the two conferences are complementary.

Further details may be obtained from Dr. H. J. Vink, Floralaan 142, Eindhoven, Holland.

Crystal data and numerical structure factor tables

The Geological Society of America, 419 West 117 Street, New York 27, New York, announces that it has reprinted its special Paper No. 33, Numerical structure factor tables (Crystal structure and X-ray diffraction) by M. J. Buerger. The volume contains 119 pages, is cloth bound, and is available at $1.50 per copy.

The Society also announces that it has in the press extensive tables of crystal data as its Memoir 60. The book will have two parts. Part I (Systematic), by W. Nowacki, gives all substances crystallizing in each of the 219 distinguishable groups. Part II (Determinative) by J. D. H. Donnay and others, permits a crystal to be identified from its cell dimensions and space group. Formula index and name index serve as a guide to the literature. The Donnay-Harker tables of space-group criteria are reprinted in an appendix. The volume is expected to have over 750 pages and will be cloth bound. The Society will make it available at only slightly above bare manufacturing cost. The price will be about $5.00. Readers who wish to be notified when the book is available should write to the Society at the above address.

Register of British manufacturers

We have received the 1954 edition of the F.B.I. register of British manufacturers. It gives details of the many British manufacturers who are members of the Federation of British Industry. The register gives information under several headings, namely, Products and services, Advertisements, Addresses; Trade associations, Trade marks, Brands and trade names. The book consists of 952 pages and, as usual, is well produced.

The register is published by Kelly's Directories Ltd., and Iliffe and Sons Ltd., Dorset House, Stamford Street, London, S.E.1. The price is 42*s.* (post free).

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions.* (Price 2*s.* 6*d.* including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any Bookseller.

Atomic energy—how it all began*

By Professor O. R. Frisch, O.B.E., F.R.S., Cavendish Laboratory, University of Cambridge

A brief review of the main steps which led to our present knowledge of atomic nuclei; some remarks on how fission might have been discovered, and how it actually was; various facts which decided the feasibility both of controlled and explosive chain reactions; and some personal reminiscences from the days when fission was a new toy, exciting and sometimes dangerous.

In so far as anything has a beginning, the story of atomic energy began in 1896. Roentgen had just discovered his X-rays, which, in his apparatus, emerged from the glass wall where it was struck by the cathode rays which he generated inside. The glass wall showed bright fluorescence, and Henry Becquerel thought that perhaps the fluorescence and the emission of X-rays were connected. So he took a variety of fluorescing compounds and placed each on a photographic plate, wrapped in black paper. None of them caused any blackening, with the exception of those which contained uranium. By testing other uranium compounds, Becquerel soon found that the fluorescence was irrelevant, the uranium content decisive. The fluorescence was a red herring which, all the same, had guided Becquerel to the discovery of radioactivity.

Though Becquerel did not know it, he had observed the first direct manifestation of atomic energy. Marie and Pierre Curie, who took up the study of the new phenomenon, did not know it either. They found that uranium ore, such as pitchblende, had a stronger blackening effect than the uranium contained in it, and they spent years of hard and tedious labour in trying to isolate the components responsible, and eventually discovered, first polonium, then radium. Also, they found that radium compounds are always a little warmer than their surroundings; 1 gram of radium generated heat at a rate of about 100 cal/h.

A gram of coal, when burnt, liberates about 8000 cal, as much as 1 gram of radium gives out in 3 days. But 1 gram of radium is not burnt out after 3 days; indeed after 25 years its energy output has dropped by only 1%. So its energy supply is many thousand times as large as that obtained from chemical reactions, and the source of that energy was a complete mystery for a while.

But not for long. Rutherford had already some evidence that radioactivity was the transformation of one element into another. This was not chemistry: it went, indeed, against the firm belief of the chemists that elements were immutable. Transformation of one element into another—and in particular of a base element into gold—had been the dream of the alchemists, a dream that had died when, in the 18th and 19th centuries, chemistry was given firm scientific foundations. But now the "philosopher's stone" had been discovered after all, and though at first only some elements could be transformed, it offered another gift, strangely analogous to the gift of eternal youth which alchemists had expected from it: the gift of virtually unlimited energy supply.

The radium atom transforms itself into an atom of radon by splitting off an atom of helium and flinging it away with great violence, at a speed of about one-tenth that of light; the heat which radium generates is really the heat of friction generated when these helium atoms are brought to rest. But how do they acquire that enormous speed? Where is the force that could impart it?

RUTHERFORD: THE ATOMIC NUCLEUS

These and similar questions were answered in 1912 when Rutherford put forward his bold idea of the atomic nucleus, a central, positively-charged core contained in each atom. It was this nucleus which embodied the permanent features of each atom; the electrons which surround it are merely a light garment, assembled to neutralize its electric charge. It is the radium *nucleus* which breaks up into a radon nucleus and one of helium, and it is the mutual electric repulsion of the two which gives the light helium nucleus its tremendous speed. Rutherford had to assume that atomic nuclei were about ten thousand times smaller in diameter than a typical atom, and his estimate has been confirmed by a variety of methods. A radon nucleus and a helium nucleus on the point of separation repel each other with a force of many pounds, and that is how a helium nucleus can acquire a speed about one-tenth that of light, even though the force acts only over an effective distance of 10^{-12} cm.

So a gram of radium contains many thousand times as much energy as a gram of coal, but it is not a useful energy source. In the first place, it is some million times more expensive than coal; secondly, the energy evolution is spread over several thousand years, and all attempts to speed it up by high temperatures or chemical agents were completely unsuccessful. That, of course, is not surprising, since the nucleus is beautifully sprung by the layers of electrons which surround it, and quite impervious to the gentle buffetings from neighbouring atoms, even at the highest temperatures. Some much more violent blow was needed to influence what goes on in a nucleus.

Once again it was Rutherford who first succeeded in causing an atomic nucleus to break up. The nucleus was one of nitrogen, and the required blow was inflicted by the impact of a helium nucleus from radium (or one of its derivatives). The energy given off in the break-up was greater than the energy of the bullet. There was great excitement that "the atom had been split," and hopes were expressed that the controlled liberation of the energy from atomic nuclei was round the corner. But the physicists knew it wasn't; for every hit on a nitrogen nucleus many thousand bullets were wasted, their energy frittered away in scattering the atomic electrons. The popular press raised another storm of excitement in 1932 when Cockcroft and Walton succeeded in breaking up lithium

* A lecture delivered in London to the Education Group of The Institute of Physics on 20 October, 1953.

nuclei by the impact of protons which they themselves had accelerated with the help of a high voltage. The newspapers told us how many times a battleship could cross the Atlantic on one ounce of lithium fuel; but all that was pointless. Millions of protons had to be accelerated to cause one hit, on an average; and every hit produced an amount of energy equivalent to that of only some hundred protons. The fact that nuclear hits could now be scored at rates many thousand times greater was a tremendous help to the scientists, and progress was speeded up enormously; but it did not turn a process of energy dissipation into one of energy production.

THE NEUTRON

But in the same year 1932, Chadwick discovered the neutron, and that discovery, though much less publicized, opened new vistas. A neutron, in passing through matter, does not waste its energy in colliding with electrons; it goes on until it hits a nucleus. Furthermore, a neutron, having no electric charge, is not repelled by the nucleus and thus can enter it even if it is quite slow; the energy which is liberated in the break-up can be millions of times greater than the energy of the neutron which causes it.

However, there is still waste in the reaction by which the neutrons are produced. Neutrons were first discovered in the break-up of beryllium when struck by a helium nucleus from a radioactive source; much greater numbers of neutrons were soon obtained with the help of particles accelerated artificially. But in either case many thousand bullets are wasted for each hit which produces a neutron.

Still, neutrons were the means by which the large scale release of energy from nuclei was eventually achieved, but it was necessary first to find a reaction in which neutrons were produced from neutrons, cutting out the wasteful behaviour of charged bullets. The first hope of such a reaction was raised quite soon after the discovery of the neutron; it was based on early measurements—later found to have been wrong—of the mass of the neutron.

Accurate measurements of the masses of nuclei are of great importance in predicting the amount of energy which can be obtained from a given nuclear reaction. This is so because of Einstein's famous equation $E = mc^2$, indicating the change m in mass which a system undergoes on receiving or giving up the energy amount E. This equation holds whatever the nature and amount of E, whether it is nuclear, chemical or mechanical. A watch when wound is heavier than when run down, by about a millionth of a millionth of a milligram, which is, of course, far too little to measure (indeed less than what is rubbed off in winding it!). A flash-light battery stores about a hundred thousand times more energy, but that still means a mass change too small to be observed. In the transformation of radium the mass loss amounts to about one part in ten thousand and thus becomes observable. But in principle there is no difference between those cases, and it is misleading to state, as the daily Press has done, that atomic energy, in contrast to chemical energy, depends on Einstein's principle of the equivalence of mass and energy. In both forms of energy production the mass of the system is changed, though in the chemical case that change is unobservably small. What Einstein's formula does is to limit the amount of energy which could possibly be obtained from a given amount of matter.

Now the masses of light nuclei had, in 1932, been measured with good accuracy with the help of mass spectrographs; but the mass of the neutron had (and still has) to be inferred from suitable nuclear reactions, with the help of Einstein's equation. At first a value was published which was a little too small, and one consequence of this was that the beryllium nucleus should be unstable, capable of disintegrating into two helium nuclei and a neutron. Actually beryllium shows no signs of spontaneous break-up, that is, of radioactivity. But perhaps, it was argued, it would break up when struck by a neutron, with the liberation of another neutron; both neutrons might strike two further beryllium nuclei, and a kind of avalanche, or chain reaction, might be started. For that chain reaction to develop, a sizable chunk of beryllium was needed, or else most of the neutrons would leave the material without hitting a beryllium nucleus. Such an experiment was actually carried out: a large amount of beryllium was assembled in a deserted spot, by remote control. Nothing happened, and I do not know if the result of the experiment was ever published, or who did it; the news of it reached me by way of rumour.

Later measurements gave a larger mass for the neutron, and it was then clear that the beryllium nucleus is lighter than two helium nuclei and one neutron. The break-up, which can indeed be caused by the impact of a sufficiently fast neutron, thus does not liberate but consumes energy. The chain reaction will come to an end after a few links at the most, when the neutrons liberated have become too slow to cause any further break-up.

THE CHAIN REACTION

It is easy to be wise after the event, and to construct the kind of argument by which a physicist, say in 1935, could have shown how to achieve a working chain reaction. Beryllium is useless, he would have said, because it is too light. All light nuclei are indeed lighter than the fragments into which they might get broken, so breaking them will not give us a source of energy. But, as one goes to heavier nuclei, one finds that the mass defect—the amount by which the mass of a nucleus is smaller than that of all the protons and neutrons of which it consists—does not grow as fast as the mass itself, and the heaviest nuclei should be capable of breaking up into lighter ones, with the liberation of large amounts of energy and probably of some neutrons. Those neutrons might stimulate the break-up of further heavy nuclei and there we would have our chain reaction.

All the facts on which this argument is based were known about 1935; why then did no physicist make it? The reason, paradoxically, is that the instability of heavy nuclei was known and thought to be well understood. The radium nucleus breaks up into a helium nucleus and a radon nucleus; since the latter is only about 2% lighter than the radium nucleus we may consider the process as the emission of a helium nucleus from a radium nucleus. The reason why a radium nucleus, on an average, waits some two thousand years before breaking up lies in the very violence with which the helium nucleus, once emitted, is pushed away by the electric field of the remainder. If it were pushed all the way from the edge of the nucleus it would acquire more energy than it actually does. So it seems that the helium nucleus begins to get pushed away from the remaining nucleus at a point a little way from its edge, having somehow ignored the electric potential gradient from the edge of the nucleus to that point. Since 1928 it has been known that quantum mechanics can account for this passage of a helium nucleus through such a "potential barrier" and indeed for many detailed features of this type of radioactivity.

In particular this theory explained why no larger fragments were observed. Energetically they would be possible; the radon nucleus, formed from radium by emission of a helium nucleus, is similarly unstable, and in fact radium is a link in a chain of transformations in which five helium nuclei are successively emitted. Why do they come out one by one? Five helium nuclei could be united into one neon nucleus, with the liberation of some extra energy; yet the emission of a neon nucleus is never observed. The quantum theory explains that: according to its formulae, the heavier neon nucleus would have vastly greater difficulty in penetrating the potential barrier, and it is much easier for the same amount of nuclear matter to seep out in the form of five successive helium nuclei than in one big lump.

What all physicists overlooked was that this argument fails for the break-up of a heavy nucleus into two approximately equal parts. In that case the available energy becomes so large that there is practically no potential barrier to be overcome. Fermi in 1934 started to study the effect of neutron impact on uranium nuclei and found evidence for the formation of several types of (radioactive) nuclei; his work was extended in Berlin and Paris, and later elsewhere. But all the scientists concerned interpreted their results under the assumption that nothing bigger than a helium nucleus could be emitted in the reaction between a uranium nucleus and a neutron. One chemist, Ida Noddack, suggested in 1934 that perhaps some of the nuclei found by Fermi belonged to much lighter elements; but the physicists thought they had good reason to disregard that suggestion. As time went on, the results got more and more complex and their interpretation more and more artificial.

NUCLEAR FISSION

The deadlock was finally broken by the two chemists, Hahn and Strassmann (Berlin), who tried to pin down one of their products as radium and found that, though chemically similar to radium, it was actually barium. That discovery, rightly rewarded by the Nobel Prize in 1945, set off a world-wide rush of activity. Lise Meitner, who for many years had shared the direction of the Radioactivity section of the Kaiser Wilhelm Institut für Chemie with Hahn, had a few months previously left Germany, and I had joined her in Sweden for Christmas when she received a letter from Hahn telling her of his result. It took her a little while to make me listen, but eventually we got to arguing about the meaning of Hahn's result, and very gradually we realized that the break-up of a uranium nucleus into two almost equal parts was a process so different from the emission of a helium, or conceivably of a neon, nucleus that it had to be pictured in quite a different way. The picture is not that of a "particle" breaking through a potential barrier, but rather the gradual deformation of the original uranium nucleus, its elongation, formation of a waist and finally separation of the two halves. The striking similarity of that picture with the process of fission by which bacteria multiply caused us to use the phrase "nuclear fission" in our first publication.

That publication was somewhat laboriously composed by long-distance telephone (Prof. Meitner had gone to Stockholm, and I had returned to Copenhagen) and eventually appeared in *Nature* in February 1939. But in the meantime other things had happened, and at bewildering speed. Bohr was in the United States when the paper by Hahn and Strassmann arrived there, and at the Washington meeting of the American Physical Society on 26 January he related it to his colleagues, together with the picture which Lise Meitner and I had formed and of which I had told him on my return from Sweden. Several experimentalists immediately went to their laboratories (some did so before Bohr had finished speaking!) in order to get experimental proof. The picture was that of two fairly large nuclei flying apart with an energy of nearly two hundred million electron volts, more than ten times the energy evolved in any other nuclear reaction, and it took only a few hours to demonstrate the large pulses of ionization caused by those fission fragments. Nor were those experimenters slow in making their results known: the daily Press reported their results within a day or two after Bohr had first drawn their attention to our interpretation of Hahn's results.

It was about a week later when Bohr was informed by a letter, which one of his sons wrote him from Copenhagen, that I had also found the large pulses of ionization some weeks before. I had not written to him myself, wanting to make sure and to follow up various questions, but I had told his son. Among the many things for which I am grateful to Bohr is the fervour with which he battled with American journalists to establish my priority on that point. It must have been in the course of that battle that I was first stated to be Bohr's son-in-law, a statement that has since got into several books although I was then unmarried, and Bohr has no daughter.

The most striking feature of this novel form of nuclear reaction was the large energy liberated. The ionization pulses could be demonstrated with surprisingly simple apparatus, a pair of headphones and one valve for amplification; and it was computed that the energy from a single fission could make a visible grain of sand make a visible jump (if we could so direct it!). But the really important question was whether neutrons were liberated in the process, and that was a point which I, for one, completely missed.

In fact Prof. Meitner suggested to me the possibility of neutron emission; but we didn't notice that this made a chain reaction possible. Later, Møller in Copenhagen pointed out this possibility to me, but I replied that the continued existence of uranium deposits showed that there were few neutrons and no chance of a chain reaction! Halban, Joliot and Kowarski in Paris, however, took the possibility of a chain reaction seriously and started measurements, and so did several teams in the United States and elsewhere. Other people must have thought of it without doing or publishing anything. According to the picture of fission the two fragments were born in a strongly deformed and hence excited state, and might each be expected to send out at least one neutron, on the average; this must have been clear to many nuclear physicists, as well as the fact that two neutrons emitted for one absorbed is amply enough for a chain reaction.

DELAYED NEUTRONS

One small effect, discovered by Roberts, Meyer and Wang early in 1939 turned out to be of enormous importance: the presence of delayed neutrons. They are emitted, not in the fission act by fragments still "hot" from the upheaval that made them, but seconds later as a consequence of the radioactive transformation of some of those fragments. It is on those neutrons that the possibility of a controlled reaction largely depends. If there were no delayed neutrons there would be an extremely sharp transition between the subcritical state (where the number of neutrons absorbed or escaping is greater than the number of neutrons produced)

and the supercritical state in which the neutrons would multiply with explosive rapidity. To get a steady power supply one would have to adjust the fraction of neutrons lost with an accuracy of about one part in a million; this would be very difficult and would be upset by the slightest disturbance, making the operation unsteady and dangerous. The delayed neutrons, although there is only one for about a hundred fissions, act as a kind of brake: the chain reaction has to wait for them if the system is only slightly supercritical, and the multiplication of the neutrons is very much slowed down, making the control of the reaction very easy.

To return to my argument about the continued existence of uranium deposits: how do they remain stable despite the emission of more than two neutrons in each fission? Well, in the first place uranium ore is not pure uranium but contains other elements as well, capable of absorbing the neutrons before they cause fission and thus quenching the chain reaction. Pure uranium was at first thought capable of sustaining a chain reaction, but later measurements showed that it is not; too many of the neutrons are absorbed by the uranium without causing fission.

Lise Meitner had shown in 1938 that neutrons can be absorbed—without causing fission—by ^{238}U which makes up the bulk of natural uranium, and she had studied various features of that absorption process. From those features Bohr concluded, early in 1939, that the fission caused by slow neutrons happened with nuclei, not of ^{238}U but of a slight admixture (0·7%) of ^{235}U. That surprising conclusion was based on rather subtle arguments and at first met with some scepticism; but it was proved correct a couple of years later when small amounts of ^{235}U had been separated out with the help of a mass spectrometer.

THE UTILIZATION OF ATOMIC ENERGY

As soon as the essential possibility of a chain reaction became clear, the thought of a devastating new weapon appeared in many people's minds. It seemed obvious that once a supercritical amount of uranium was assembled the first neutron that came to enter it would make it explode with terrible force. Bohr thought otherwise. From the scanty data then available he concluded that a mass of pure uranium, however large, would never explode, there being too much absorption without fission; this conclusion was fully confirmed by later measurements. The only way to get a chain reaction, Bohr thought, was to combine uranium with some light element which would slow down the neutrons and thus bring the small admixture of ^{235}U into play, with its preference for slow neutrons; and such a combination, even if it were made to explode, would react so slowly that it would be no worse than gunpowder, blowing itself apart before more than a tiny fraction of the uranium had suffered fission. In this he was partly right: that was how Fermi and others produced the first chain reaction on earth on 6 December, 1942, in Chicago. Unhappily, Bohr was wrong in thinking that it was the only way. He had not reckoned with the ingenuity and tenacity of physicists and engineers, both British and American, driven by the fear that the Germans might be first in producing the decisive weapon. Within a few years they developed a laboratory instrument, the mass spectrometer, into a vast industrial enterprise, capable of separating many pounds of ^{235}U from tons of natural uranium. Freed from the neutron-absorbing ^{238}U, the ^{235}U can undergo a chain reaction with unslowed neutrons, multiplying at a tremendous rate and resulting in a devastating explosion.

But the wartime demand for a decisive weapon was also a great stimulus for the development of controlled chain reactors, or "piles" as they came to be called, because the first one consisted of uranium rods and graphite blocks, piled up in regular array; a very inappropriate name for the alternative construction of a tank containing uranium rods immersed in heavy water, used first in Canada and later in France and Norway. This was so because it was soon realized (though I do not know who was the first to do so) that the capture of neutrons in ^{238}U might lead to the formation of an element—plutonium—which could serve as a nuclear explosive like ^{235}U but could be separated from uranium much more easily, by chemical methods.

Atomic energy will soon be harnessed to serve industry and commerce; but I like to remember the days when chain reactors were a new toy for physicists to play with. For a few exciting weeks we in Los Alamos were allowed to study the growing amounts of ^{235}U which were being prepared for the first atomic bomb. The material was made into little bricks for us, and by piling those bricks into assemblies of various shapes, surrounded by various materials, we could check the calculations of critical size, a nearby counter clicking faster and faster as one brick after another was added until we could sit back and watch the counting rate rise exponentially. We knew that we could stop the chain reaction at any time by removing a few bricks (or a few blocks of the surrounding material), and we never felt that those experiments were dangerous; actually two people died by causing, through some minor slip, a large quantity of neutrons to pass through their bodies, and I once nearly became the third.

There is much more one could say about the development of atomic energy, but I merely said I would tell you how it all began, and this lecture must have an end somewhere.

The physical properties of gelatin solutions and gels

By A. G. Ward, M.A., F.Inst.P., The British Gelatine and Glue Research Association, London, N.7

The composition and structure of gelatin are related to the method of manufacture, and to the physical properties of gelatin solutions and gels. The importance of the isoionic point, molecular weight distribution and a gel-forming factor are discussed in connexion with the experimental results.

In the past five years there has been a considerable revival of interest in the properties of the gelatin/water system. In earlier investigations, especially in the period 1920 to the early 1930's, it was not possible to interpret experimental results in terms of structure. Modern techniques of polymer and protein physics and chemistry permit a new approach to these problems, with an enhanced chance of successful theoretical analysis.

The present review will first outline those aspects of the manufacture and chemical structure of gelatin which are of significance in considering the physical properties. The lack of information on these subjects has reduced the value of many investigations. The remainder of the paper will include accounts of the more important recent investigations of physical properties.

THE PRODUCTION OF GELATIN FROM COLLAGEN

Commercially there are two processes for the production of gelatine* from collagen. In addition there are a wide variety of collagenous tissues which are used as raw material for either process. The most important raw materials are skin, bone and sinew, but they can be derived from a number of animals—cow, sheep, pig, etc. Skin is mainly obtained from tannery wastes, and will have received a short treatment with alkali in the tannery. Freshly flayed and refrigerated pigskin is supplied by the meat-packing factories in the U.S.A. to the gelatine manufacturer, and is an important raw material. Sun-dried bone from India and Pakistan is the best type of bone for gelatin manufacture. The mineral portion is removed by means of hydrochloric acid, leaving "ossein" which has a very high content of bone collagen. Sinew is only of minor importance as a raw material, but it is of interest in being a different tissue. Although very different in size and appearance, all these raw materials are largely made up of a collagen type protein, and may be treated by similar methods to give gelatine.

The two principal manufacturing processes have as their aim the extraction of the gelatine at as low a temperature as possible, so that the gel-forming power of the gelatin obtained, and also its solution viscosity, are not seriously reduced. In the first method,[1] the "alkali process," a prolonged treatment of the raw material with a lime suspension (pH 12–12·5) at 10–20° C enables the gelatin to be extracted by subsequent warming with water at or near neutrality (pH 5–8). In addition, a number of impurities present in skin, bone and sinew dissolve in the lime suspension and so do not appear in the final gelatine solution. The majority of British commercial gelatines are prepared by this process.

The second or "acid" process does not involve a pretreatment except to reduce the pH for extraction to the region of pH 4. At this pH, extraction is quite rapid even at moderate temperatures. Certain of the impurities are insoluble at this pH. A substantial proportion of U.S. gelatines are prepared by this process from pigskin.

* The term "gelatine" is used to mean a commercial or laboratory product inclusive of ash and other impurities. The term "gelatin" is restricted to the derived protein obtained from collagen, irrespective of whether in gelatine or animal glue.

It is important for any person working with gelatin to know by which of these two processes the gelatin was made, since the two types differ in properties. Fortunately, it is fairly simple to detect to which type the gelatin belongs (p. 87) since their isoionic points are widely different.

Apart from differences arising from these two processes, commercial gelatines vary widely in physical properties, according to the history of the raw material, the heating used in their extraction, and the care used in subsequent processing. It is therefore necessary to specify most carefully the gelatine used in any investigation. Much of the early work has little or no value owing to neglect of this precaution.

THE STRUCTURE OF THE GELATIN MOLECULE

Gelatin is derived from the protein collagen. It is likely that, in common with the parent protein, it consists of polypeptide chains. Investigations on skin collagen[2] suggest that the chains in the unmodified material are either very long indeed, or are cyclic, since no evidence can be found of the existence of the amino group which would be expected at the chain end.

Gelatin (Fig. 1) has readily detectable amino end groups which can be estimated, and, for the higher grades of gelatin,

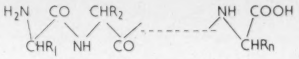


Fig. 1. The gelatin chain molecule. R_1, R_2 the different side groups of the constituent amino acids

these are in the proportion of one for every 60 000 molecular weight.[3] The simplest assumption is that the gelatin molecule is composed of a single chain terminating in an amino group at one end, and a carboxyl group at the other. The molecular weight varies from chain to chain over a wide range but averages about 60 000 for a high-grade product. This view is confirmed by the results of Pouradier and Venet who give figures in this region for the number average molecular weight M_n by osmotic methods.[4,5,6] If the gelatin molecule contained two or more chains then M_n, by end group methods, would be very much smaller than M_n by osmotic methods.

The gelatin chain is built up from the linking together of some eighteen types of amino acids. It is unlikely, owing to the complexity of the problem, that much progress will be made in the near future in determining the *order* of these amino acids along the gelatin chains, but the *relative proportions* are fairly definitely known. The figures, based on those of Tristram,[7] are given in Table 1.

The gelatin molecule presents to the exterior environment the active groups listed in Table 1. Of the active groups, some (carboxyl, imidazole, amino, guanidino) can ionize in aqueous solution. The range of pH over which this occurs is related to the pK values given in Table 1. At pH 2, the carboxyl groups are all unionized and so the gelatin molecules will carry their maximum positive charge, due mainly to the ionized amino and guanidino groups. As the pH is raised, the carboxyl groups begin to ionize and at pH 6·5 they all

carry negative charges. At higher pH's the positively charged groups begin to lose their charge, imidazole by pH 8, ε-amino by pH 11·5, and guanidino at very high pH. At the isoionic point (which, in the absence of salts, will also be the isoelectric point) the net charge is zero, the number of positive and negative charges being equal.

Table 1. *Amino acid composition of gelatin (after Tristram*[7])

Amino acid	*Number of groups per 60 000 mol. wt.*	*Active groups*	*pK (40° C)*	*Sign of charge*
Glycine	215	None	—	—
Alanine	63	None	—	—
Valine	17	None	—	—
Leucine	16	None	—	—
Isoleucine	8	None	—	—
Methionine	4	None	—	—
Threonine	11	Hydroxyl	—	—
Serine	18	Hydroxyl	—	—
Proline	77	None	—	—
Hydroxy-proline	66	Hydroxyl	—	—
Phenyl alanine	9	None	—	—
Tyrosine	3·5	Phenolic hydroxyl	10	Negative
Histidine	3	Imidazole	7	Positive
Lysine	19	ε-amino	10	Positive
Hydroxylysine	4·5	ε-amino	9	Positive
Arginine	29·5	Guanidino	> 12	Positive
Aspartic acid	30	Carboxyl	3·5–5·0	Negative
Glutamic acid	46	Carboxyl		
Amide*	28	—	—	—
Total hydroxyl	103			
Total basic	56	Alkali processed gelatin (omitting the tyrosine OH group)		
Total acidic	76			
Total basic	56	Acid process gelatin (omitting the tyrosine OH group)		
Total acidic	48			
Total peptide bonds	639			

* Present in acid process gelatin. The actual figure is taken from collagen.

In native collagen some 37% of the carboxyl groups are present as amide groups ($-CONH_2$) which do not ionize at any pH. The amide groups are not destroyed by the commercial "acid" process for gelatine manufacture (although *prolonged* treatment with acid converts them to carboxyl groups), so that the gelatin molecules made in this way have more ionizable basic than acidic groups (Table 1). The isoionic point is then at about pH 9·0 to 9·4. Alkali processed gelatin, in which the majority of the amide groups have been converted to carboxyl groups by the pretreatment, have isoionic and isoelectric points about pH 5·1 to 4·8 depending on the thoroughness of the pretreatment.

The position of the isoionic point of a gelatin sample will have an important influence on its properties, since it indicates the sign of the charge which will be carried at other pH's. At pH 6 an alkali process gelatin will carry a negative charge, whereas an acid process gelatin will be positively charged.

The importance of the non-ionizing groups must not be neglected. Of these the most numerous are the hydroxyl groups derived from serine, threonine, hydroxy-proline, hydroxy-lysine and tyrosine, and the peptide groups –CO–NH– in the chain backbone. The peptide group can readily form hydrogen bonds, and in solid gelatin or collagen, the need to form the maximum number of these bonds helps to determine the structure.[8, 9]

No significant difference in molecular structure has yet been demonstrated between gelatins derived from different tissues or from different mammalian species. There is, however, a slight difference in composition between certain soluble collagens (often termed procollagen) and the insoluble collagen. This difference would be reflected in the gelatin derived from them. There is also a difference of amino acid composition between fish collagens and mammalian collagens, particularly in the hydroxy amino acids.[10,11] This may correspond to the different behaviour of the "gelatin" made from fish collagen, which appears to have a lower gel-forming power.

The structure of gelatin may then be summarized: a gelatin sample will contain polypeptide chains of widely differing length, terminated at one end by an amino group and at the other by a carboxyl group, with other active groupings distributed along the chain. The two variable structural properties which would be expected to influence the physical behaviour are:

(i) the distribution of molecular weights, which will be reflected in different values for the mean molecular weights (M_n, M_w, etc.);
(ii) the isoionic point, and consequently the charges carried at different pH values.

The measurement of these structural features will be briefly considered before examining how far they serve to explain the known behaviour of gelatin.

Molecular weight and molecular weight distribution. Fractionation into portions of reasonably uniform molecular weight is an essential preliminary to molecular weight measurement for materials in which a wide range occurs in any one sample. The solubility of most macromolecules depends on the molecular weight, and early work using salts or alcohol as precipitants has shown that gelatin may be fractionated in this way. Two methods have appeared recently using this basic principle in rather different ways. In the first,[4] the conditions are adjusted so that the less soluble portions of the gelatin are thrown out of solution, by addition of alcohol, not as a precipitate, but as finely divided droplets of a very viscous solution or "coacervate." The drops settle to give a lower layer which can be separated from the bulk of the solution. The fractions obtained may be refractionated by a repetition of the process. This method has been developed by Pouradier and Venet for the preparation of fractionated material to use in other investigations, rather than for establishing the molecular weight distribution in the original samples. The range of values of molecular weight occurring can, however, be established.

The second method[12] involves salting out of solution the complex formed between gelatin and sodium dodecyl sulphate (S.D.S.). The formation of this complex has been very fully investigated.[13,14] It would appear that dodecyl sulphate ions become bound by their sulphate "heads" to the positively charged groups and also to the –CO–NH– groups of the gelatin backbone chain, leaving their hydrocarbon "tails" to the outside. The net effect is to give the gelatin molecule a surface which is much more hydrocarbon in character than gelatin itself, so that addition of salt (NaCl) in the range from 0·4 *M* to 1·0 *M* enables successive fractions to be brought out of solution. The physical form in which the less soluble portion separates is again a coacervate. This probably has advantages in securing sharper fractionation than precipitation of a solid. The gelatin in the complex is separated

from the S.D.S. by precipitation from 2 : 1 acetone/water, in which the S.D.S. is left as a solution.

The latter method enables recoveries of nearly 100% to be obtained, giving some idea of the distribution of molecular weights. A recent modification by Stainsby[15] of the alcohol method will enable distributions to be determined by this method also.

It should be emphasized that neither method is likely to give very sharp cuts of the molecular weight distribution so that the fractions must not be regarded as completely homogeneous. The properties of fractions obtained by the two methods are essentially concordant, which suggests that the fractionations are on the basis of molecular weight, since the molecular surfaces are so different in the two procedures.

The only molecular weight determinations for these fractionated materials have been made by Pouradier and Venet, using an osmotic pressure method. For many practical purposes it is of value to relate the molecular weights to dilute solution viscosity measurements. This is complicated for gelatin (p. 89) by the substantial variations of dilute solution viscosity with pH and salt content of the solution. Pouradier and Venet have chosen to carry out the viscosity measurements at the isoionic point, in the absence of salt, extrapolating the reduced viscosity $\eta(\text{specific})/C$ to infinite dilution to give an "intrinsic" viscosity, $[\eta]$. Although this has certain theoretical advantages, the value obtained is very sensitive to small errors both in the pH used and in the amount of salt present. This is avoided by Stainsby, Saunders and Ward, who use pH 7 in the presence of 1 *M* NaCl. Under these conditions, small variations of pH and salt concentration do not change the value of $1/C \log_e \eta_{rel}$, the reduced viscosity (see Fig. 5).

Pouradier and Venet have shown that the values of M_n, determined by osmotic pressure, and the intrinsic viscosities, for fractions of different gelatins falls on a single line when plotted, provided that the isoionic points of the gelatins are the same.[6] The relation $[\eta] = KM^{\gamma}$ was found to hold, with different constants for the two types of gelatin, acid and alkali processed. The actual equations are:

(*a*) Using results for fractions of:

two gelatins pI's = 4·75 and 4·80
$[\eta] = 1{\cdot}66 \times 10^{-5} M^{0{\cdot}885}$

(*b*) one gelatin pI = 8·8
$[\eta] = 1{\cdot}10 \times 10^{-4} M^{0{\cdot}74}$

Pouradier and Venet's results show that the range of molecular weight occurring in high-grade gelatin samples is from 20 000 to 150 000. The true range is rather wider than this since there is some inhomogeneity, especially of the high molecular weight fractions.

The isoionic and isoelectric points. Janus, Kenchington and Ward have developed a very rapid method for determining the isoionic point.[16] The gelatin solution is passed through a mixed bed of anionic and cationic ion exchange resins (Amberlite IR 120 and IRA 400 by Rohm and Haas Co.). By this means, the salts present are reduced to so low a figure that the pH of the solution coincides with the isoionic point, at least for solutions of concentration not less than 1%. This method is also useful for the rapid preparation of deionized gelatin.

Previous methods have used the existence of a maximum or minimum for certain properties measured as a function of pH—viscosity, swelling, opalescence of the dilute gel. The direct measurement of the isoelectric point involves finding the pH of zero mobility for gelatin molecules in an electric field. The results obtained by the deionization method are in agreement with the previous methods, within the errors of pH measurement. In addition it would not appear that salts such as sodium chloride displace appreciably the isoelectric point. A few typical results are given in Table 2.[16]

Table 2. *Isoionic points for several gelatin samples*

		Isoionic (or isoelectric) point			
Precursor	*Process*	*A*	*B*	*C*	*D*
Hide	Alkali	5·08	5·10	5·09	5·06
Hide	Alkali	4·84	4·84	4·87	—
Ossein (bone collagen)	Alkali	5·05	5·04	5·06	—
Sinew	Alkali	4·85	4·87	4·84	4·82
Pigskin	Acid	9·3	9·3	—	—
Ossein	Acid	8·7	—	—	—

Methods: *A*, deionization; *B*, maximum turbidity (ash free); *C*, maximum turbidity (with salt); *D*, minimum viscosity.

It is of interest to note that the two limiting values, for acid extracted gelatin from untreated collagen, and for gelatin from collagen receiving prolonged alkali pretreatment, do not appear to depend on the tissue or species from which the collagen is taken. This suggests, in confirmation of the analytical results, that the proportions of the ionizable groups, and of amide groups, must be very similar in all types of mammalian collagen.

Having described the preparation of gelatin fractions, the determination of molecular weights, and of isoionic (or isoelectric) points, it now remains to see to what extent the properties of gelatin solutions, gels and films may be explained in terms of these structural properties.

CHEMICAL COMPOSITION OF FRACTIONS

Pouradier, Roman and Venet[17] have shown that, within rather large limits of experimental error, the nitrogen content, the isoelectric point and the amount of hydrogen ions required to discharge the carboxyl groups, does not vary from fraction to fraction for a single gelatin. This agrees with expectation, if the low molecular weight materials are merely shorter fragments of chains similar to those from which the long fragments are derived.

This view is also confirmed by the work of Courts,[3] who shows that the same terminal amino acids, carrying free α-amino groups, appear when gelatin is degraded, as are found in the original material (aspartic and glutamic acids, serine, threonine, alanine and especially glycine). This suggests that, in forming gelatin, the collagen chains have been broken at certain weaker links, which are still fairly common in gelatin itself, and break during further hydrolysis.

VISCOSITY AND RIGIDITY

The rheological behaviour of gelatin is very temperature dependent and three regions may be distinguished. The boundaries between the regions depend on the gelatin used and also on concentration, but the figures given here refer to high-grade gelatins at concentrations up to 5–6%. Above 35° C, the gelatin solutions do not show abnormalities of flow and a viscosity coefficient may be satisfactorily measured. Difficulties may arise at higher temperatures, or at acid or alkaline pH's, due to chain breaking, so that in accurate measurements it is necessary to avoid these conditions. Below 20° C, except at very low concentrations, the solution will exist as a gel. The properties of the gel depend not only on the temperature of measurement but on the thermal history. These and other aspects have been very fully

reviewed by Ferry[18] recently and will not be further considered here. Between 20 and 35° C, the gelatin solution may exist either as a gel, as a viscoelastic liquid or as a liquid showing anomalous viscosity, according to the temperature and thermal history.

It appears likely that over 35° C the gelatin molecules in solution are discrete. Although in concentrated solutions they may interfere with each others' motion, they do not link together for appreciable times. In the gel, rigidity is obtained, once sufficient linking points have been formed, to make a three-dimensional network. With time, however, further molecules become alined together so that the minute network fibrils are sufficiently large and organized to give an X-ray diffraction picture.

Although these general conceptions are unlikely to be changed, little is known concerning the points of the gelatin molecule which are capable of forming the cross links. The gel-inhibiting effect of substances known to break hydrogen bonds,[18] such as urea, would suggest that hydrogen bonding is involved. In the region between gel and viscous liquid (20–30° C) more or less aggregation of molecules will have occurred, the anomalies in behaviour increasing with the size of the aggregates; until the complete network is set up and the gel formed.

Measurements by Stainsby and Saunders[19] have shown that the viscosity of fractions of gelatin, when measured at 40° C and 5·5% concentration, is, at least to a fair approximation, a unique function of the dilute solution viscosity. This applies irrespective of the isoionic point of the gelatin, and confirms the absence of aggregates in these conditions. No such relationship holds between the rigidities of the fractions and the dilute solution viscosity even for gelatins of the same isoionic point. It appears clear that some—at present unknown—structural feature of the gelatin molecule is involved in gel formation, and that for undegraded gelatins it is of much greater importance than molecular weight.

A new method[20] for rigidity determination has proved very useful in this connexion. The gelatin solution is set as a cylindrical column, in precision bore tubing, and matured under appropriate conditions. When air pressure P is applied at one end of the column, the lower surface displaces a volume Q given by $PR^4/8LG$, where L is the column length, R the radius and G the modulus of rigidity. The tube is connected at its lower end to a capillary tube of radius r carrying an etched scale, the intervening space and part of the capillary being filled with mercury. The volume change Q therefore becomes a displacement h of the mercury in the capillary, and G is given by $PR^4/8Lr^2h$. Since h can be made about 6 cm, the measurement becomes both simple and accurate. By setting gels at 0° C, measurements may be made at this temperature. The temperature can then be raised to 10, 17·5 and 25° C in turn, the rigidities being measured after allowing 3 h at each temperature.

Fig. 2 shows results obtained by this method for fractions of a single gelatin, similar sets of results having been obtained for other gelatins. The rigidity at 0 and 10° C is clearly independent of the reduced viscosity and hence of the molecular weight. There will be a lower limit of molecular weight beyond which this constancy no longer holds, and recent work suggests that this may correspond to a reduced viscosity of about 0·2.[21] At higher temperatures or at low concentration the molecular weight exerts a greater influence on the values of the rigidity obtained.

The results conflict with the conclusions of Ferry.[22] Ferry made use of unfractionated materials of which the molecular weights had been previously determined by Scatchard, Oncley, Williams and Brown.[23] The gelatins included a series in which a single gelatin had been subjected to different degrees of controlled degradation, giving number average molecular weights between 45 000 and 17 000. The rigidities were measured by the method of transverse vibrations,[24] the modulus obtained being independent of frequency over a twofold range. Ferry and co-workers have since shown[25] that the dynamic moduli obtained by this method are in excellent agreement with static results with a concentric cylinder apparatus. Ferry obtained[22] an empirical relation between the rigidity at a given temperature and M_w, the results being illustrated in Fig. 3. These experimental results can possibly be reconciled with the behaviour of gelatin fractions, if it is considered that the factor responsible for gel formation is slowly destroyed by heating and that, near neutrality, the rate of destruction must be similar to the rate of fall of molecular weight. The dependence of the rigidity on molecular weight obtained by Ferry then becomes fortuitous, resulting from the use of heat degraded samples.

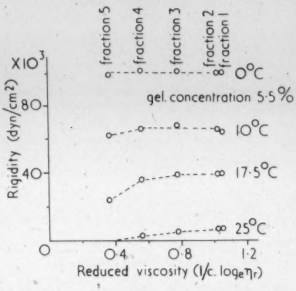


Fig. 2. Rigidity of 5·5% gelatin gels, prepared from fractionated gelatin, as a function of reduced viscosity.

(Precursor: alkali processed hide)

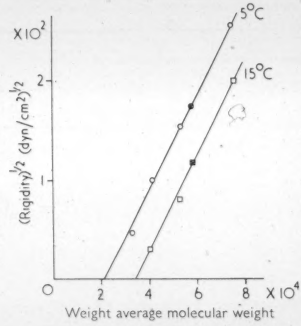


Fig. 3. (Rigidity)$^{\frac{1}{2}}$ as a function of weight average molecular weight for a series of degraded gelatins. (After Ferry[22])

Open points, degraded samples; crossed points, mixture.

The emergence of a third structural feature, which, together with the isoionic point and molecular weight distribution, controls the behaviour of gelatin, explains much of the confusion experienced by both users of gelatin and those carrying out research on it, since this feature will influence other properties also.

This has been clearly shown by Pouradier and Venet,[6] and also by Stainsby and Saunders[19] using gelatins of the same isoelectric point. The gel melting point for fractions of different gelatins does not depend only on molecular weight

or reduced viscosity. This is not very surprising owing to the close connexion between gel formation and the melting point.

A further property where the results cannot be explained by molecular weight and isoionic point differences is the threshold concentration of alcohol at which turbidity is observed in a gelatin solution, under carefully standardized conditions. Fig. 4, based on the data of Pouradier and

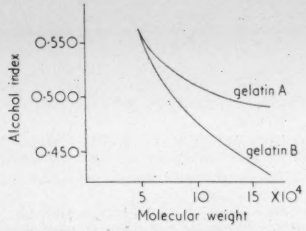


Fig. 4. Relative concentration of alcohol to cause turbidity in isoelectric solutions (gelatin concentration 1% before alcohol). Gelatin *A*, alkali treated ossein precursor, pI = 4·8. Gelatin *B*, alkali treated skin precursor, pI = 4·75. (After Pouradier and Venet[6])

Venet,[6] shows the results for two gelatins of closely similar isoionic point, which give quite different curves for the required alcohol concentration as a function of molecular weight.

It remains to be seen which properties of gelatin depend on these structural features, but investigations are still at an early stage.

THE FLEXIBILITY OF A GELATIN MOLECULE

Direct evidence concerning the flexibility of the gelatin molecule is difficult to obtain. It is possible, however, to use the force exerted by the charges carried on the molecule to test its extensibility. Stainsby[26] has measured the dilute solution viscosity of gelatin as a function of pH over the range 0·5 to 12·5, by adding hydrochloric acid or sodium hydroxide to a salt free gelatin solution. Fig. 5 shows the result obtained. A qualitative explanation can be given, if the reduced viscosity

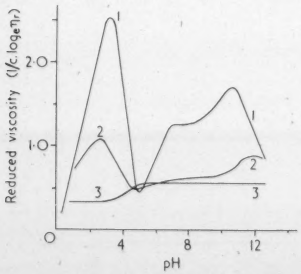


Fig. 5. Reduced viscosity of gelatin in dilute solution as a function of pH and salt content. Alkali treated precursor, pI = 5·08

Curve 1, 0·2% ash-free gelatin in pure water; curve 2, 0·2% ash-free gelatin in 0·017 *M* salt; curve 3, 0·2% ash-free gelatin in 1·00 *M* salt.

is regarded as a measure of the axial ratio of an ellipsoid bounding the gelatin molecule. At pH 5·1, the isoionic point of the gelatin used, equal positive and negative charges are distributed along the molecular chain causing the maximum folding and contraction. As the pH is raised or lowered, a net negative or positive charge develops, so that the repulsive forces produced cause some uncoiling of the molecule. By pH 3·1 or pH 10·7 this uncoiling has become very marked. Further additions of alkali or acid, while increasing the net charge, reduce the resultant forces owing to the action of the counter-ion atmospheres. Maxima are therefore shown. In more concentrated solutions, where the existence of the maxima has been observed by many workers, the effect is not so well marked since the greater additions of acid and alkali needed to change the pH give increased ion concentrations which reduce the forces.

Additions of neutral salts also reduce the effective forces between the charged sites, so that the maxima are less marked. At the isoionic point, the decrease in forces causes a slight stretching. At high salt concentrations the maxima are completely suppressed.

These results are in close accord with those obtained by Katchalsky[27] and co-workers for amphoteric polyelectrolytes.

STRESS RELAXATION

Miller, Ferry, Schremp and Eldridge[25] have carried out an interesting investigation into stress relaxation in gelatin gels, using unfractionated material. The measurements were made in a concentric cylinder apparatus in which the outer cylinder could be rotated rapidly through an angle and fixed in position. The torque on the inner cylinder was measured by a very stiff torsion support so that the angular movement of the inner cylinder was almost negligible compared to the initial movement of the outer cylinder. By measuring the torque as a function of time, the stress relaxation at constant strain was obtained. A typical result is shown in Fig. 6.

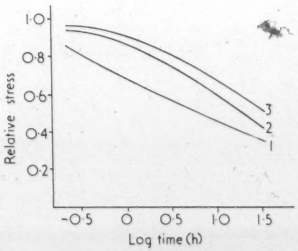


Fig. 6. Relative stress plotted against log (time). Alkali treated ossein precursor. (After Miller and others[25])

Curve 1, gel aged at 16·2° C for 1·1 h; curve 2, gel aged at 16·2° C for 12·0 h; curve 3, gel aged at 16·2° C for 24·3 h.

The identity of the dynamic moduli and the static moduli as ordinarily measured, demonstrates the absence of relaxation times from 10^{-3} to 10^{2} sec. The observed relaxation requires relaxation times from 10^{3} to 10^{6} sec at about 10–15° C, with shorter times at higher temperatures. The instantaneous rigidity is not changed by the relaxation process. It would appear that, during stress relaxation, gelatin molecules detach themselves from stressed positions and re-attach themselves

in an unstressed position. The new bonds formed are then similar in type to those originally present, and so give rise to similar short time moduli.

It is probable that the gelatin gel is not a static structure, but one in which, even in the absence of stress, solution and redeposition is occurring. The increase in rigidity of gels,[18] when matured at constant temperature, would arise from such rearrangements.

CONFIGURATION OF THE MOLECULE IN THE SOL AND GEL STATE

One of the most characteristic features of the behaviour of collagen is the thermal contraction which occurs when it is heated in water to a temperature of 60–70° C. The thermally contracted material has lost the typical X-ray diffraction picture of collagen, and it would appear that a new molecular configuration has replaced the low temperature structure. In the normal structure of collagen it is usually considered[8] that hydrogen bonds are formed *between* the chains. Thermal contraction may then correspond to the formation of hydrogen bonds *within* the chains in place of those stable at lower temperatures. Simultaneous stretching and cooling of thermally shrunk collagen will partially restore it to the previous condition.

Since a stretched dried gelatin gel can also show thermal shrinkage[28] it is interesting to see whether the configuration of gelatin in the gel state differs from that in the sol.

A re-examination has been carried out by Robinson and Bott[29] of the mutarotation effect previously investigated by Smith,[30] Kraemer and Fanselow,[31] and Ferry and Eldridge.[32] Robinson and Bott made use of results for synthetic polypeptides which showed a clear connexion between the optical rotation and the extent to which the polypeptides chain was present in the α configuration (intra-molecular hydrogen bonds) or the β configuration (inter-molecular hydrogen bonds). This suggested that the mutarotation effect, in which the optical rotation changes as gelatin is cooled from sol to gel, might be attributed to a similar cause. Films prepared by drying gelatin from the gel, and from the sol (e.g. at 55° C) were shown to have markedly different optical rotations, confirming results of Smith.[30] In addition the two types of dried film have been shown to have different solubility properties,[33] the hot dried film showing much greater cold water solubility than the film obtained by drying the gel.

Bradbury and Martin[28] have examined films of gelatin prepared by drying at 20° C and at 60° C. Tensile strength measurements show the cold films to be substantially stronger, but if the films are conditioned to $>75\%$ R.H. before breaking, the "hot" prepared films showed much greater extensibility. X-ray examination shows crystallinity and orientation in the films formed by drying the gel, whereas the "hot" dried films give amorphous diagrams. Bradbury and Martin interpret their results in terms of the higher crystallinity of the low temperature films, but it may well be that both the nature of the chain configuration and the more macroscopic structural features play a part. Jopling[34] has thrown further light on properties of the two types of film by examining their swelling behaviour.

CONCLUSIONS

Developments in research on gelatin have reached a stage where important advances are likely in the next few years. Of primary interest must be the mechanism of gel formation, where the isolation and identification of the unknown factor remains as the outstanding problem.

ACKNOWLEDGEMENT

I am indebted to the Council of the British Gelatine and Glue Research Association for permission to publish this review.

REFERENCES

(1) AMES, W. M. *J. Soc. Leather Trades Chem.*, **33**, p. 407 (1949).
(2) BOWES, J. H. *Nature* [*London*], **168**, p. 514 (1951).
(3) COURTS, A., in WARD, A. G. *Nature* [*London*], **171**, p. 1099 (1953).
(4) POURADIER, J., and VENET, A. M. *J. Chim. Phys.*, **47**, p. 11 (1950).
(5) POURADIER, J., and VENET, A. M. *J. Chim. Phys.*, **47**, p. 391 (1950).
(6) POURADIER, J., and VENET, A. M. *J. Chim. Phys.*, **49**, p. 85 (1952).
(7) TRISTRAM, G. R. *The Proteins*, Vol. 1A (New York: Academic Press, Inc., 1953).
(8) AMBROSE, E. J., and ELLIOTT, A. *Proc. Roy. Soc., A*, **205**, p. 47 (1951).
(9) PAULING, L., and COREY, R. B. *Proc. Nat. Acad. Sci., U.S.A.*, **37**, p. 272 (1951).
(10) NEUMAN, R. E. *Arch. Biochem.*, **24**, p. 289 (1949).
(11) NEUMAN, R. E., and LOGAN, M. A. *J. Biol. Chem.*, **154**, p. 314 (1950).
(12) STAINSBY, G., SAUNDERS, P. R., and WARD, A. G. *J. Polymer Sci.*, **12**, p. 325 (1954).
(13) PANKHURST, K. G. A. *Surface Chemistry*, p. 109 (London: Butterworths Scientific Publications Ltd., 1949).
(14) PANKHURST, K. G. A., and SMITH, R. C. M. *Trans Faraday Soc.*, **40**, p. 565 (1944); **41**, p. 630 (1945).
(15) STAINSBY, G. Private communication.
(16) JANUS, J. W., KENCHINGTON, A. W., and WARD, A. G. *Research*, **4**, p. 247 (1951).
(17) POURADIER, J., ROMAN, J., and VENET, A. M. *J. Chim. Phys.*, **47**, p. 887 (1950).
(18) FERRY, J. D. *Advances in Protein Chemistry*, **4**, p. 1 (1948).
(19) STAINSBY, G., and SAUNDERS, P. R., in Ref. (3) above.
(20) SAUNDERS, P. R., and WARD, A. G. *Proc. 2nd Int. Congr. on Rheology* (London: Butterworths Scientific Publications Ltd., 1954).
(21) STAINSBY, G. *Adhesives and Resins*, **1**, p. 222 (1953).
(22) FERRY, J. D. *J. Amer. Chem. Soc.*, **70**, p. 2244 (1948).
(23) SCATCHARD, G., ONCLEY, J. L., WILLIAMS, J. W., and BROWN, A. *J. Amer. Chem. Soc.*, **66**, p. 1980 (1944).
(24) FERRY, J. D. *Rev. Sci. Instrum.*, **12**, p. 79 (1941).
(25) MILLER, M., FERRY, J. D., SCHREMP, F. W., and ELDRIDGE, J. E. *J. Phys. Coll. Chem.*, **55**, p. 1387 (1951).
(26) STAINSBY, G. *Nature* [*London*], **169**, p. 662 (1952).
(27) KATCHALSKY, A. *J. Polymer Sci.*, **7**, p. 393 (1951).
(28) BRADBURY, E., and MARTIN, C. *Proc. Roy. Soc., A*, **214**, p. 183 (1952).
(29) ROBINSON, C., and BOTT, M. J. *Nature* [*London*], **168**, p. 325 (1951).
(30) SMITH, C. R. *J. Amer. Chem. Soc.*, **41**, p. 1135 (1919).
(31) KRAEMER, E. O., and FANSELOW, J. R. *J. Phys. Chem.*, **29**, p. 1169 (1925).
(32) FERRY, J. D., and ELDRIDGE, J. E. *J. Phys. Coll. Chem.*, **53**, p. 184 (1949).
(33) PINOIR, R., and POURADIER, J. *C.R. Acad. Sci.* [*Paris*], **227**, p. 190 (1948).
(34) JOPLING, D. W. *Research*, **6**, p. 27S (1953).

Heat transfer by free convection in an open thermosyphon tube

By B. W. MARTIN, M.Sc., A.M.I.Mech.E., A.F.R.Ae.S., and H. COHEN, M.A., Ph.D., A.M.I.Mech.E., A.F.R.Ae.S.,
King's College, University of Durham, Newcastle upon Tyne

[*Paper first received* 21 *July and in final form* 11 *November*, 1953]

An account is given of an experimental investigation into the heat transfer by free convection in a heated vertical tube closed at its lower end. To cover a wide range of physical properties of the convective fluid, glycerine, water and air were used. The trends are shown to be in generally close agreement with theoretical prediction.

The proposal to cool the blades of high temperature gas turbines by the flow of a fluid contained in radial cylindrical cavities within the blades has brought into focus a method of transferring heat concerning which there is little reliable knowledge at the present time. This free convection mode of cooling was first suggested by E. Schmidt[1], and the blades of the turbine built to his specification contained such cavities, all sealed near the blade tip and each communicating with a reservoir of cool fluid in the turbine rotor. The centrifugal force field created by rotation caused the heated fluid adjacent to the cavity wall to move towards the centre of rotation. This was simultaneously replaced by an outward radial flow of cooler fluid from the reservoir along the core of the cavity. Some warming of the entering fluid in the reservoir was inevitable since it was surrounded by the heated annulus of fluid leaving the cavity. Although the cavities were very small, the high rate of rotation and the consequent strength of the centrifugal field was capable of producing high rates of heat transfer from the turbine blade.

It is clear that this system is a special case of a natural convection heat transporting arrangement in which the cool fluid is totally surrounded by the hot fluid and where there is no dividing metallic wall so that the fluid itself has to create its own internal boundary. Experimental investigation of this system presents a difficult problem when an attempt is made to operate under accelerations comparable with those occurring in an actual turbine. However, an expedient approach to the problem is the use of a scaled-up static heated vertical cavity in which the centrifugal acceleration in the turbine is replaced by the acceleration due to gravity.

The effect of the confining walls of the tube, coupled with the presence of a counter flow, therefore gives rise to a fundamental problem in the study of laminar and turbulent free convection. The use of a static rig has other advantages, since it permits the measurement of heat transfer rates at relatively low Grashof numbers, which is a desirable preliminary before work at high Grashof numbers is attempted.

DESCRIPTION OF EXPERIMENTAL APPARATUS

The experimental apparatus used is shown diagrammatically in Fig. 1, and consists basically of a heated vertical tube, closed at its lower end, while the upper end opens into a suitably mounted large reservoir. A cooling coil is built into the reservoir, so that the heat conveyed to the latter by heated fluid emerging from the tube is carried away by the cold water circulating through the coil.

The brass tube used has an internal diameter of 2 in. and an overall length of 47·5 in. There are in all five electrical heating coils, each formed by winding special resistance cable supplied by Pyrotenax Ltd. on the spirally-grooved outer surface of the tube, afterwards applying a heavy thickness of solder, so that the coil is completely encased. Around the tube is a 6 in. diameter copper outer casing, $\frac{1}{4}$ in. thick. The annulus between tube and casing is filled with granulated cork which serves as an insulating material, and minimizes external heat losses. Five copper-constantan thermocouple junctions are embedded in the inner wall of the tube, at positions along the length corresponding to those of the five heaters. There are also three thermocouple junctions in the outer casing so that the temperature difference between tube and casing may be ascertained and an estimate made of the external heat losses. The temperature of the cold fluid entering the tube is measured by means of a probe thermocouple junction placed on the axis of the tube at the orifice.

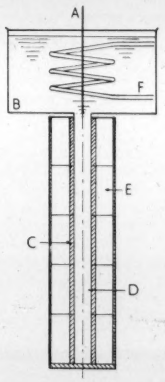


Fig. 1. Diagrammatic, sectioned view of apparatus
A, probe thermocouple; *B*, tank; *C*, heater element; *D*, tube; *E*, granulated cork.

Tests have been carried out using chemically pure glycerine, air, and water as the working fluids. Throughout the tests, the procedure was the same. It consisted of applying a potential difference across the output terminals of the Variac transformer, providing power for the heaters. Adjustments were then made to rheostats in series with the five heating elements so that when steady conditions obtained, the temperature of the tube was practically uniform over its

length. Absolute uniformity of temperature could be guaranteed only by the use of an infinitely large number of heaters. This, of course, was not a practical possibility and the findings detailed below are therefore approximate to that extent. In view of the relatively high thermal conductivity of the brass tube, it is improbable that any substantial errors are introduced by the use of a finite number of heaters. The temperature difference between the entering fluid, as recorded by the probe thermocouple, and the wall temperature of the tube was regarded as the effective temperature difference promoting heat transfer. The net heat input was obtained by subtracting the external heat loss, estimated from the temperature difference between tube surface and the outer casing, from the gross heat input. For this purpose, a separate closed end test had previously been carried out, in which the tube cavity was filled with granulated cork.

EXPERIMENTAL RESULTS AND DISCUSSION

Notation

α, coefficient of cubical expansion of the fluid;
a, internal radius of tube;
C, specific heat of the fluid;
g, external acceleration (gravity in this case);
k, thermal conductivity of the fluid;
l, length of tube;
μ, absolute viscosity of the fluid;
ρ, density of the fluid;
Q, total rate of heat transfer from tube;
ΔT, temperature difference between entering fluid and tube wall;
$\kappa = k/C\rho$, thermal diffusivity of the fluid;
$\nu = \mu/\rho$, kinematic viscosity of the fluid;
Nu, Nusselt number;
Gr, Grashof number;
Pr, Prandtl number;
t, Grashof–Prandtl product divided by length: radius ratio of tube.

The ten quantities first named in the notation may be regarded as comprising those which are likely to exert an appreciable influence on the problem of heat transfer by natural convection in a tube. Some of these quantities are interdependent, and their number may be reduced by replacing thermal conductivity by thermal diffusivity, and absolute viscosity by the kinematic viscosity of the fluid. These parameters are defined above.

The remaining variables may be reduced to manageable proportions by dimensional analysis. It has been shown[2,3] that probably the most convenient and significant dimensionless groups thus obtainable are

(*a*) $\dfrac{l}{a}$ — Length radius ratio of tube.

(*b*) $\dfrac{\nu}{\kappa} = Pr$ — Prandtl number.

(*c*) $\dfrac{\alpha g a^3 \Delta T}{\nu^2} = Gr$ — Grashof number or non-dimensional temperature difference using tube radius.

(*d*) $\dfrac{Q}{2\pi l k \Delta T} = Nu$ — Nusselt number or non-dimensional heat transfer rate based on tube radius.

Experience has shown that the relation between these four groups may be expressed by an equation of the form:

$$Nu = f\left(Gr \,.\, Pr \,.\, \frac{l}{a}\right) \text{ where } t = \frac{Gr \times Pr}{l/a}$$

The nature of the function remains to be determined. It is, however, clear that for the present case in which l/a is constant at 47·5

$$Nu = F(Gr \,.\, Pr)$$

The product of the Grashof and Prandtl numbers is sometimes referred to as the Rayleigh number, or modified Grashof number. Long experience has established that it is more appropriate to free convection flow than the conventional Grashof number. The property values required to express temperature difference and heat transfer rate in their dimensionless forms are applied throughout at the tube wall temperature. Calculations suggest that for the above purpose this temperature is likely to be the most practicable approximation to a representative fluid mean temperature. The numerical values of the dimensionless parameters thus obtained from analysis of the test results are given in the appendix.

The results of the first test, using glycerine, are shown graphically in Fig. 2, both the Nusselt and Grashof–Prandtl numbers being plotted to logarithmic scales, and covering a

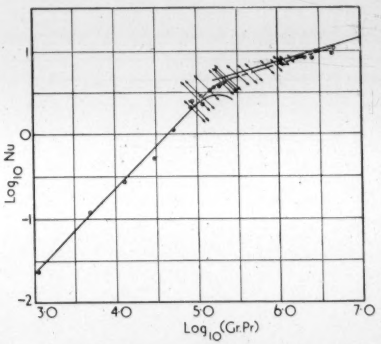


Fig. 2. Comparison of theoretical and experimental laminar flow regimes

—— theoretical laminar flow for $Pr = \infty$.
○ experimental glycerine test.
⟷ flow oscillation.

range such that $3 < \log_{10}(Gr \,.\, Pr) < 6{\cdot}65$. The corresponding range of Prandtl numbers is from 7600 to 66, covering a tube temperature variation from 60 to 200° F. The lower part of the curve is a straight line of unit slope for Grashof–Prandtl numbers between 10^3 and $10^{5{\cdot}2}$ implying a linear relationship between the two variables plotted. On analysis, the equation takes the form

$$Nu = \frac{Gr \,.\, Pr}{42\,400}, \text{ or } Nu = \frac{t}{892}$$

At a value of $Gr \,.\, Pr$ of 10^3 only the top heater is effective in conveying heat to the fluid, but as the temperature difference is increased, the lower heaters are successively brought into

use, until at a value estimated by interpolation at $10^{4\cdot25}$, with a corresponding Nusselt number of 0·42, all heaters are effective. Apparently convection begins at the orifice, and the effective length of tube is extended towards the closed end with increasing Grashof–Prandtl product, until the whole mass of fluid is involved. Prior to this condition, an isothermal stagnant region must exist at the lower end, wherein the fluid is at tube temperature.

Beyond a value of $Gr \cdot Pr$ of $10^{5\cdot2}$ the straight line bends over and terminates at $\log_{10} Gr \cdot Pr$ equal to 5·34, when the Nusselt number is 4·32. If the Grashof–Prandtl product is further increased, it becomes impossible to avoid a regular periodic oscillation of the fluid entry temperature, which persists up to a value of $Gr \cdot Pr$ equal to $10^{6\cdot15}$. Once established, it may be preserved by careful experimentation down to a value of $10^{4\cdot95}$. Visual inspection of the flow at the orifice showed a regular periodic discharge into the reservoir of heated glycerine adjacent to the tube wall, the discharge. period corresponding to a falling fluid entry temperature. This oscillation has a maximum intensity, in terms of temperature change, at a value of $Gr \cdot Pr$ of $10^{5\cdot50}$ and it will be seen from Fig. 2 that if a reasonable mean curve is drawn through the lines of oscillation, its slope is approximately 0·28 over the greater part of the oscillatory range. The power law relating the dimensionless groups in this region appears to be

$$Nu = \frac{(Gr \cdot Pr)^{0\cdot28}}{6\cdot92} = \frac{t^{0\cdot28}}{2\cdot35}$$

Further study of the oscillatory range showed that the period of oscillation varied in an inverse manner with heat input (Fig. 3). A typical waveform shown in Fig. 4 suggests

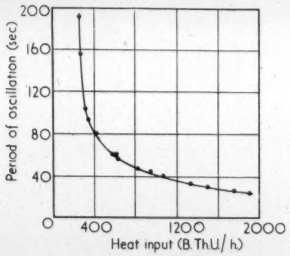


Fig. 3. Period of oscillation against heat input

that the oscillation is of a non-linear type, the time for the temperature to rise being much greater than that needed for it to fall. Having regard to the high viscosity of glycerine, and the correspondingly large Prandtl numbers, it seems probable that the entire test took place under laminar flow conditions. This is borne out by the index value of 0·28 in a region where the limiting case, for a tube of infinitely large radius, is of laminar boundary layer flow up a heated vertical plate. This limiting case has been studied theoretically by Polhausen[4] and Squire[5] and found to give an index of 0·25.

Lighthill's recent work[6] on free convection in tubes, based on purely theoretical considerations, shows close agreement for laminar flow with the experimental curve of Fig. 2. In the range of linear variation, the theoretical and experimental curves are coincident if an infinite Prandtl number is assumed. A similarity regime is predicted up to the condition where the full length of tube is utilized, envisaging a stagnant pool of fluid near the closed end until the similarity solution obtains, when the Nusselt number is predicted to be 0·414. This also is in accord with experimental evidence. The similarity regime implies that the distribution of velocity and temperature across each cross-section of the tube is similar, but that the scale of variation

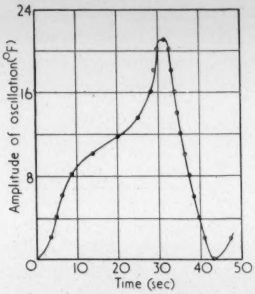


Fig. 4. Typical oscillation waveform

increases linearly with length moving from the bottom of the tube towards the top. Beyond this condition, the flow is supposed to fill the tube without similarity, leading to crimped similarity profiles near the closed end, and the theoretical line ultimately curves over to a maximum Nusselt number of about 3 and terminates at a slightly lower value. The tendency to a maximum appears experimentally, but the two curves diverge and regular oscillations commence before the maximum is clearly attained.

The third regime of laminar flow which is forecast is of a boundary layer type which does not fill the tube. For an infinite Prandtl number, the graphical interpretation of this regime is an almost straight line having a mean slope of about 0·27 which tends towards 0·25 in the limiting case of a vertical plate. This lies slightly above the mean experimental line passing through the oscillatory zone.

A most interesting feature is that the two branches of the theoretical curve, representing boundary layer and non-similarity flow, never meet, and that the method of transition is not mathematically resolved. It is suggested that transition is resolved in practice by the appearance of the oscillations previously described, which indicate the periodic formation and decay of a boundary layer. The establishment of a boundary layer, with generally higher fluid velocities, leads to a more rapid intake of fluid to the tube, and since the time is reduced for preheating the entering fluid just above the orifice, the heating effect is less and the entry temperature falls. The decay of the boundary layer and the resumption of non-similarity flow, for which the velocities are comparatively low, gives rise to enhanced preheating, and so the entry temperature slowly rises, the cycle of events being indefinitely repeated. The disappearance of oscillations at high Grashof–Prandtl numbers coincides with the establishment of a stable boundary layer but since this does not occur until $10^{6\cdot15}$ the transition range between the two types of flow appears to be remarkably large.

The results of tests carried out with air and water are shown in Fig. 5. In the case of air, compression of the Grashof–Prandtl range is almost entirely due to the inverse

relation of the product term $\alpha/\kappa\nu$ with temperature, and the small Nusselt numbers merely reflect the very low heat transfer rates that could be achieved. The variation of Prandtl number during the test was between 0·71 and 0·69, and the greatest tube temperature was 270° F.

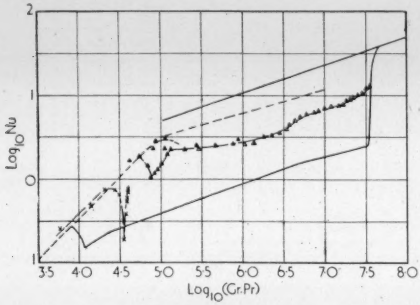


Fig. 5. Comparison of theoretical and experimental turbulent flow regimes

— — theoretical laminar flow for $Pr = 1$.
—— theoretical turbulent flow for $Pr = 1$.
△ experimental water test.
× experimental air test.

During the early part of the test, where $3{\cdot}77 < \log_{10} Gr \,.\, Pr < 4{\cdot}25$, it may from previous knowledge be inferred that similarity and non-similarity laminar flow are present. The slope of the curve is near to unity, but with respect to the corresponding curve for glycerine, it is slightly displaced to the left. This is in agreement with Lighthill's theoretical conclusion, and is apparently due to the difference in Prandtl number. It will be seen that the curve then bends sharply to a maximum Nusselt number of 0·75 well below the predicted turning point for laminar flow, and forms a crevasse at $\log_{10} Gr \,.\, Pr$ equal to 4·55 with a minimum value of Nu of 0·19. The dimensionless heat transfer rate then rises steeply at an almost constant Grashof–Prandtl number.

Having regard to the nature of the working fluid, with a low viscosity, the evidence points to a transition to turbulent flow at $\log_{10} Gr \,.\, Pr$ equal to 4·3. Such a transition is forecast by Lighthill but, for a Prandtl number of unity and length-radius ratio of 50, the breakaway from laminar flow is expected at $\log_{10} Gr \,.\, Pr$ equal to about 4·0. Although the predicted crevasse is less steep than that obtained with air, the lowest Nusselt number of 0·19 is in fair agreement with the theoretical estimate of 0·25. The reduction in heat transfer rates due to the incidence of turbulence is explained by supposing that the flow is seriously impeded by turbulence while the turbulent flow fills the tube without similarity, and more than offsets the increased effectiveness of the flow in transferring heat which turbulence would otherwise yield. This, indeed, appears to be the case. When a turbulent boundary layer is established, at much higher Grashof–Prandtl numbers, the reverse will be true. This last regime was obtained by an extrapolation of Saunders' results.[3]

The results obtained with water, where the Prandtl number varied from 8·5 to 1·8, tend to confirm the general conclusions drawn from the air test. Breakaway from laminar flow seems to occur at a higher Grashof–Prandtl number, about $10^{4{\cdot}75}$, a little below the suspected oscillatory region. The crevasse is not nearly so steep as before, and its depth is less, the least Nusselt number being about unity. At $\log_{10} Gr \,.\, Pr$ equal to about 5·1, the heat transfer curve becomes almost flat and remains so until a value of 6·0, when it ascends in a series of curved steps, yielding a Nusselt number of 13·64 at a Grashof–Prandtl number of $10^{7{\cdot}55}$. At this point the tube temperature was 214° F, and since the system was not pressurized, the water boiled.

Throughout the test the dimensionless heat transfer rates are well below those predicted for both laminar and turbulent boundary layer flow, and it seems clear that this is due to the impeding effect of turbulence before a boundary layer is set up. On the other hand, the effect with water is not so severe as predicted for the length-radius ratio used, with unit Prandtl number. Even for air, with a lower Prandtl number, it appears unlikely that the reduction would be as great as estimated from the theoretical analysis. The distribution of local heat transfer rates along the tube length during this test was quite different from that when using glycerine. The largest heat input was at the upper end of the tube, the intensity falling off very rapidly with length until a negligible input was required of the fourth heater. The fifth and bottom heater supplied a very moderate quantity. Such a distribution is envisaged in the theoretical consideration of turbulence, and it is suggested that laminar flow may be present at the closed end.

CONCLUSIONS

The available information suggests that turbulent flow in a closed ended tube is less effective than laminar boundary layer flow so long as turbulence interferes with the movement of the heated fluid filling the tube. Only if conditions permit the existence of a turbulent boundary layer is it likely that turbulent flow will show superior rates of heat transfer, and it may be necessary to employ a rotating rig to determine when this condition is likely to occur. Future work may also make it possible to derive some sort of experimental criterion which will indicate the onset of turbulence.

ACKNOWLEDGEMENTS

The authors wish to thank Professor M. J. Lighthill of the University of Manchester for permission to draw on information from his paper prior to publication, and they are indebted to Professor A. F. Burstall, of King's College, University of Durham for advice in connexion with the experimental work and for the provision of the necessary facilities. They would also like to thank Dr. S. Astin of Thomas Hedley and Co. Ltd., Newcastle upon Tyne for arranging the gift of glycerine and providing information regarding its physical properties.

REFERENCES

(1) SCHMIDT, E. *General Discussion on Heat Transfer. Section IV.* (London: Institution of Mechanical Engineers, 1951.)
(2) MCADAMS, W. H. *Heat Transmission.* 2nd edition. (New York: McGraw-Hill Book Co., 1942.)
(3) SAUNDERS, O. A. *Proc. Roy. Soc.*, **157**, pp. 278–91 (1936).
(4) SCHMIDT, E., and BECKMAN, W. *Tech. Mech. Thermod.*, **I**, pp. 341–9 and pp. 391–406 (1930).
(5) GOLDSTEIN, S. (ed.). *Modern Developments in Fluid Dynamics.* (London: Oxford University Press, 1938.)
(6) LIGHTHILL, M. J. Theoretical considerations on free convection in tubes. *Quart. J. Math. Appl. Mech.* To be published.

APPENDIX

Tables of numerical results

Bracketed readings denote temperature oscillations

Glycerine				*Air*				*Water*			
$Gr \times 10^{-3}$	$Pr \times 10^{-3}$	$Gr \cdot Pr \times 10^{-5}$	Nu	$Gr \times 10^{-4}$	Pr	$Gr \cdot Pr \times 10^{-4}$	Nu	$Gr \times 10^{-5}$	Pr	$Gr \cdot Pr \times 10^{-5}$	Nu
0·000 145	7·6	0·011 0	0·02	0·834	0·707	0·59	0·25	0·050 2	8·16	0·409	1·75
0·000 938	5·25	0·049 2	0·12	1·98	0·706	1·40	0·48	0·085 5	8·12	0·694	1·47
0·003 50	3·72	0·130	0·28	2·96	0·704	2·08	0·74	0·092 7	8·12	0·753	1·04
0·012 1	2·47	0·298	0·53	3·95	0·703	2·77	0·73	0·093 9	8·48	0·796	2·38
0·031 9	1·62	0·517	1·14	4·65	0·700	3·25	0·64	0·101	8·51	0·862	1·25
0·075 3	1·12	0·844	2·13	5·14	0·700	3·60	0·32	0·106	8·23	0·875	2·95
0·128 0	0·910	1·16	2·37	5·16	0·698	3·60	0·19	0·112	8·22	0·917	2·13
0·080 0	1·10	0·880	2·52	5·45	0·695	3·79	0·38	0·113	8·27	0·933	1·32
0·059 7		0·711	2·58	5·63	0·692	3·90	0·55	0·127	8·11	1·03	1·48
0·109	1·19	1·30	1·41	5·66	0·689	3·90	0·68	0·139	8·05	1·12	3·02
0·076 0		0·821	2·55	5·85	0·687	4·02	0·75	0·139	8·13	1·13	2·43
0·127	1·08	1·37	1·53	5·74	0·685	3·93	0·58	0·150	7·99	1·20	2·08
0·070 2		0·652	4·29					0·160	8·05	1·29	2·48
0·156	0·930	1·45	1·93					0·244	7·86	1·92	2·35
0·086 3		0·742	4·52					0·335	8·23	2·76	2·67
0·194	0·860	1·67	2·01					0·393	7·26	2·85	2·46
0·118	0·950	1·12	2·95					0·591	7·86	4·64	2·60
0·123		0·920	5·18					1·06	7·17	7·57	2·81
0·253	0·750	1·90	2·52					1·34	6·47	8·66	3·29
0·186	0·775	1·44	3·49					1·50	7·09	10·6	2·72
0·218	0·728	1·59	3·69					1·99	6·80	13·5	2·91
0·246		1·41	5·96					2·74	6·42	17·6	3·46
0·493	0·572	2·82	2·98					3·57	6·27	22·4	3·50
0·282	0·666	1·88	3·81					4·72	6·00	28·3	3·80
0·320	0·610	1·95	4·12					6·39	5·41	34·6	4·24
0·304	0·532	1·62	6·49					7·46	5·03	37·5	4·91
0·598		3·18	3·33					9·18	4·79	43·9	5·52
0·378	0·572	2·16	4·32					10·9	4·57	50	5·96
0·291		1·55	6·39					14·1	4·28	60·5	6·36
0·589	0·532	3·14	3·16					19	3·93	74·6	6·74
0·349		1·71	6·87					24·1	3·55	85·5	7·24
0·726	0·490	3·55	3·32					27·8	3·53	98·1	7·67
0·567		2·38	6·99					36·6	3·23	118	7·88
1·06	0·420	4·45	3·73					50	2·98	149	7·90
0·851		3·16	7·00					58·6	2·82	165	8·24
1·49	0·372	5·55	3·96					71·4	2·55	182	9·42
1·25		3·89	7·25					81·5	2·48	202	9·76
2·08	0·312	6·50	4·33					99·3	2·33	231	9·98
2·46		6·04	7·73					126	2·06	260	10·84
3·71	0·246	9·14	5·11					149	1·97	294	11·45
4·06		8·12	7·99					172	1·88	324	11·87
5·75	0·200	11·5	5·63					179	1·85	332	13·02
4·74		8·81	8·81					196	1·81	355	13·64
6·35	0·186	11·18	6·61								
5·35		9·51	8·65								
6·92	0·178	12·3	6·66								
6·17		10·3	8·86								
8·22	0·167	13·7	6·72								
9·19		13·4	8·57								
11·0	0·146	16·1	7·13								
13·3	0·127	16·9	8·48								
16·0	0·118	18·9	8·66								
18·9	0·112	21·2	8·47								
25·7	0·100	25·7	8·29								
32·2	0·088	28·3	9·97								
68·3	0·066	45·1	9·06								

An evaporated carbon replica technique for use with the electron microscope and its application to the study of photographic grains

By D. E. BRADLEY, Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks.

[*Paper received* 2 *October*, 1953]

A method for the production of carbon replicas by evaporation on to the specimen is described. The specimen is subsequently removed by solution. The method has been applied to the examination of photographic grains, which could not be resolved in the optical microscope or observed directly in the electron microscope. A resolution better than 100 Å is obtained.

The production of thin carbon films for specimen supports in the electron microscope was discussed in a previous paper.[1] The same technique has been applied with success to the production of carbon replicas. The method differs from that used by König[2] in that the carbon is deposited directly by evaporation. This is of some practical importance, since it is not necessary to introduce solvent vapours into a vacuum chamber, thus contaminating the pumping plant and possibly affecting the specimen. It is not yet possible to assess fully the relative merits of the new technique and the well-established silica technique,[3] though the carbon method appears somewhat simpler; the evaporation is very quick and easy to control and the carbon film is clearly visible throughout the process due to its high optical absorption properties. This is a great practical advantage. Once the apparatus has been set up, it is possible to make a replica in less than fifteen minutes. The resolution obtained is better than 100 Å.

The method has been applied to the examination of photographic grains, which cannot be viewed directly in the electron microscope without being considerably modified by the electron beam.

METHOD

The photographic emulsion to be examined was digested from the surface of the photographic plate with hot water and the suspension centrifuged at about 3500 r.p.m. for fifteen minutes. The supernatant liquid was decanted and the particles suspended in hot distilled water. A repetition of the centrifuging process removed all gelatine and other soluble contaminants. A small quantity of the aqueous suspension of the particles was dried on a clean microscope slide and carbon was evaporated normally on to the surface in a vacuum.

The specimen was then removed from the vacuum chamber and dipped in a 7% solution of Bedacryl 122 (a resin obtainable from Imperial Chemical Industries Ltd.) in benzene to form a strong backing for the carbon replica. When dry the composite film was scored into squares and floated off on to the surface of a solution of sodium thiosulphate to remove the halide particles. This operation was best performed by teasing up the corner of a square with a needle and gripping it in a pair of forceps. Both slide and forceps were then immersed in the solution and the backed replica was easily pulled away from the slide and brought to the surface. The squares were washed by floating in a bath of distilled water.

To mount the replica, an electron microscope specimen grid was slightly bent and placed on the top of a $\frac{1}{8}$ in. diameter metal peg. It was thus supported at the edges only, so as to minimize the possibility of surface tension breaking the replica. The squares of film were lifted on a metal ring tool and mounted on the grid in the order: grid—carbon—Bedacryl. A few drops of methylated spirits applied to the specimen after mounting on the grid assisted in adhering the carbon replica to the matt surface of the grid. After drying, the backing was washed away by allowing 3 or 4 c.c. of a mixture of 50% ether and acetone to flow over the grid from a burette. When dry the replica was ready for examination in the microscope.

RESULTS

Fig. 1 shows a replica of Ilford photomechanical grains. The cubic crystals are all oriented so that one flat face of each is parallel to the glass slide on which they were originally deposited. The carbon, as in the silica method, is deposited in a uniform layer all round each crystal. Contrast is produced by the differential thickness resulting from the varying shape of the film. The uniform film nature of the replica is better demonstrated by Fig. 3, showing a replica curled over at an angle to the electron beam. In this case the original particles, which were taken from an Ilford nuclear emulsion, are spherical. In Fig. 1 it is interesting to note that each grain replica contains a small dense agglomeration. It is possible that this is a particle of silver produced by "printing out" due to exposure to light during the replica process. It is tempting to suggest that these agglomerates indicate the positions of the sensitive specks in the grains. Fig. 2 shows a similar replica shadowed with gold-palladium. Some fine structure is now apparent on the grains and on the background film. This is probably due to residue left over from the centrifuging. The thickness of the carbon film can be estimated as less than 100 Å from the width of the lines indicating the walls of the small spheres appearing in Fig. 1.

It has generally been supposed that evaporated silica molecules flow round objects due to surface migration though it has been thought possible that reflexion from adjacent walls of the apparatus or gas scattering may also contribute. Carbon, while exhibiting the same effect, does not form uniform films if evaporated on to a surface in large quantities. Fig. 4 demonstrates this. The electron micrograph shows a carbon replica of polystyrene latex particles on to which carbon has been evaporated at an angle of $\tan^{-1}\frac{1}{2}$. The carbon has formed a film covering the particles and the background, though detail in the sharp shadows shows that in the area of the shadows a film of limited thickness is present. This suggests that a flow of carbon into the shadow region occurs. It is interesting to note a build-up of 200 Å of carbon on the side away from the source, also suggesting a flowing of carbon atoms, in this case over the surface of the latex particles. A build-up of 600 Å on the side nearest the source indicates that the flowing is only partial.

CONCLUSIONS

The carbon replica technique described shows promise of giving a resolution superior to that obtained by more usual methods. It is simple to carry out and, like the silica method, is most suitable for stereoscopy. Attempts are now being made to apply the method to metallurgy.

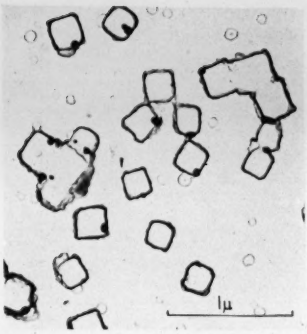


Fig. 1. Carbon replica of photomechanical particles not shadowed

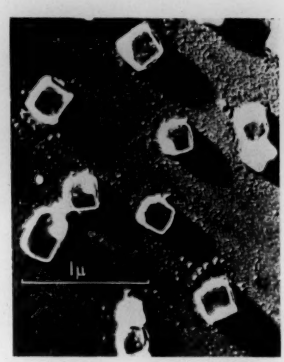


Fig. 2. Carbon replica of photomechanical particles shadowed with gold-palladium at $\tan^{-1}\frac{1}{2}$

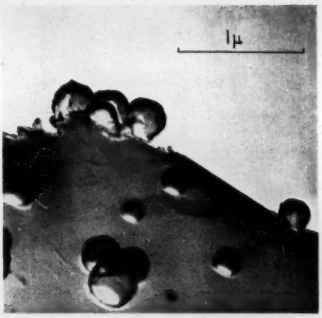


Fig. 3. Unshadowed carbon replica of nuclear emulsion particles bent over at an angle to the electron beam

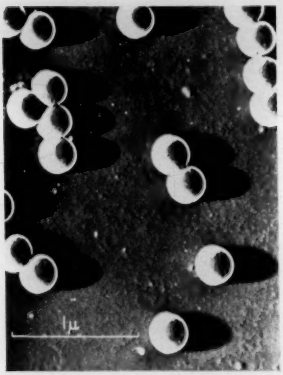


Fig. 4. Unshadowed carbon replica of polystyrene latex particles made by evaporating carbon on to the particles at $\tan^{-1}\frac{1}{2}$

ACKNOWLEDGEMENTS

The author wishes to thank Ilford Ltd. for supplying the emulsions, Mr. M. E. Haine for help and encouragement, and Dr. T. E. Allibone, for permission to publish this article.

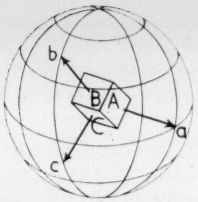
REFERENCES

(1) BRADLEY, D. E. *Brit. J. Appl. Phys.*, **5**, p. 65 (1954).
(2) KÖNIG, H., and HELWIG, G. *Z. Phys.*, **129**, pp. 491–503 (1951).
(3) HEIDENREICH, R. D., and PECK, V. G. *J. Appl. Phys.*, **14**, p. 23 (1943).

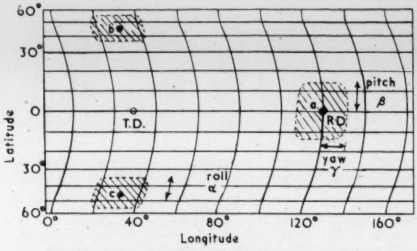
The determination of the texture of sheet steel from torque curves

By L. R. Blake, B.Sc., Ph.D., A.M.I.E.E.

See pages 99–104

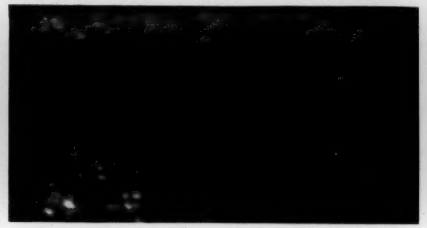


(*a*) Points *a*, *b*, *c* give the latitude and longitude of the normals to the faces *A*, *B*, *C* of the crystal.



(*b*) Co-ordinate grid to interpret the texture maps.

(*c*) Texture map of a specimen (26) with well-defined texture.



(*d*) Texture map of a specimen (57) with poor texture.

Fig. 8. Texture maps of 3·5% silicon-iron specimens

Some observations on the time edge effect in Catalin 800

By A. R. Morris, M.Sc.

See pages 104–106

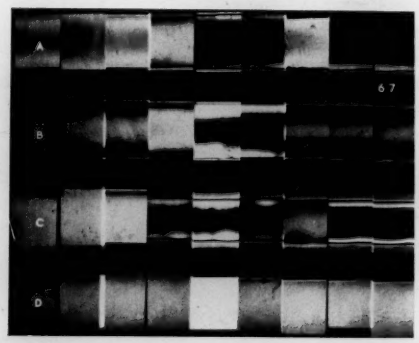


Fig. 4. A photograph similar to that from which the data illustrated in Fig. 1 was obtained

Reading from left to right there is an interval of one day between one photograph and the next.

The determination of the texture of sheet steel from torque curves

By L. R. Blake, B.Sc., Ph.D., A.M.I.E.E., Research Laboratory, The British Thomson-Houston Co. Ltd., Rugby

[*Paper first received* 6 *May and in final form* 9 *October*, 1953]

The excellent magnetic properties of cold-reduced silicon-iron strip are achieved by a close control of the mechanical and heat treatments, the object of which is to attain a high degree of preferred orientation in which the plane of the sheet is a {110} crystallographic plane and the direction of rolling is a ⟨100⟩ or cube edge direction. The torque curve is used extensively to check if the ideal orientation is achieved.

In this paper a rapid method of deducing the orientation from torque curves is proposed, applicable to the particular case of cold-reduced silicon-iron sheet. Used in conjunction with X-ray data, the method assists the more accurate assessment of the deviation from the ideal orientation; used alone it enables the departure to be measured in terms of the "root mean square angle of spread," and gives some information of the component angles of spread in different directions. The results show that the torque curve can be used with confidence as a means of control during production, and useful information of the orientation can be obtained from the curve without the need for elaborate analysis.

INTRODUCTION

The orientation of any particular crystal in a polycrystalline material can be specified by the angles α, β and γ by which it differs from some standard position.[1] If there are few crystals, then the complete specification of the orientation requires that the orientation of each crystal be given; if the number of crystals is large, the complete specification is best given by a distribution function $\Phi(\alpha, \beta, \gamma)$, where $\Phi(\alpha, \beta, \gamma)\, d\alpha\, d\beta\, d\gamma$ gives the proportion of crystals having an orientation between α, β, γ and $\alpha + d\alpha, \beta + d\beta, \gamma + d\gamma$. The information of the texture which can be deduced from a single torque curve will, in general, be very limited and inexact, and will rarely approach this ideal specification. On the other hand, it is possible by X-ray analysis to obtain a very good idea of the function $\phi(\alpha, \beta, \gamma)$, particularly with the "texture mapping camera" developed by Milner and James.[2]

Although with polycrystalline material the usefulness of the torque curve is limited and is no substitute for X-ray methods of deducing texture, there are occasions when the torque curve is of value. If in a manufacturing process a certain orientation is aimed for, the torque curve immediately shows whether it has been achieved, and the extent of the deviation of the actual torque curve from the ideal gives a rough indication of the disorientation of the material. The torque curve is also useful in conjunction with X-ray analysis when it assists the quantitative description of the texture.

Various methods of expressing the texture of polycrystalline material as deduced from a torque curve in a restricted yet helpful form have been suggested. Akulov and Bruchatov[3] tried to describe the texture by assuming only the following orientations to be present:

(100) [100]; (100) [110]; (110) [100]; (110) [110].

They assumed that m% of the material had one of these orientations, n% had another, and the remainder were random. Brailsford[4] endeavoured to describe the texture in a similar way. Bitter[1] assumed the material to have a basic orientation which is known, with deviations from this orientation in three mutually perpendicular directions. De Barr[5] did much the same. The assumptions made by Bitter and De Barr are particularly appropriate to the case of cold-rolled silicon-iron sheet and so will be considered in more detail later.

To obtain the maximum information of the texture of polycrystalline material from a torque curve, it is desirable that the values of the magnetic anisotropy constants K_1 and K_2 should be known for the material. It is worthwhile, therefore, to summarize the values of K_1 and K_2, obtained mainly with single crystals of silicon-iron, in the hope that the same values will apply to polycrystalline material of the same composition.

The values of K_1 *and* K_2 *for silicon-iron alloys.* In material with a cubic structure in which the cube edge directions are the easy directions of magnetization, and the cube diagonals the difficult directions, it can be shown[6] that K_1 is always positive and K_2/K_1 lies between $-9/4$ and $+\infty$. This is the only theoretical restriction on the values of K_1 and K_2; the actual values appear to depend mainly on the composition.

The value of K_1 for iron-silicon alloys is given in Fig. 1.[7,8] Unfortunately, the values of K_1 as obtained by various workers

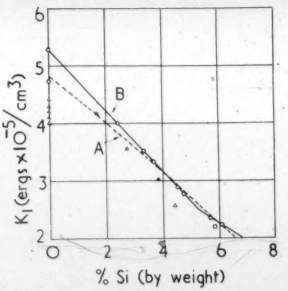


Fig. 1. The value of the crystal anisotropy constant K_1 for iron-silicon alloys

B. Tarasov's curve.
A. Bozorth's curve using same data.

differ appreciably, especially for pure iron where the variation is ±14%. The most systematic investigation was that of Tarasov, but two of his specimens of pure iron gave values for K_1 differing by 11%, and were in any case much higher than the values of others. Tarasov assumed the lower value to be in error for reasons unknown, since, as can be seen from Fig. 1, the larger value when plotted fell on a straight line through his other points. Bozorth,[8] apparently in trying to compensate for the generally lower values of other workers, assumed the curve to be lower than that of Tarasov. It appears that accuracy of K_1 to within ±5% cannot be expected.

Published values of K_2 vary widely and cannot be related to composition. When estimating K_2 from (110) disks, great accuracy is necessary since the terms involving K_2 are small in comparison with the K_1 term. Williams's[9] estimate of $1{\cdot}0 \times 10^5$ erg.cm^{-3} for K_2 is probably the most reliable figure available for 3·85% silicon-iron, but the totally different values[10] of other workers for similar compositions make it impossible to be certain of the value of K_2 for any particular specimen. Values of K_2/K_1 range from $-0{\cdot}7$ to $+0{\cdot}5$. Due to the low values of the terms involving K_2 in low silicon-iron alloys, it is normally possible to ignore them when deducing the orientation of single crystal disks, except for orientations near (111).

The value of the applied field and the use of a disk specimen. Errors in assessing the orientation can arise due to the torque being measured at too low a value of the field. The effect of H on the general shape of the torque curve for the particular case of a (110) single crystal disk can be seen from Fig. 2.

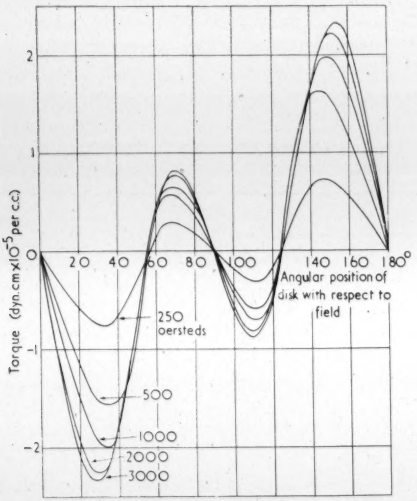


Fig. 2. The effect of the applied field on the torque curve of an essentially (110) single crystal disk

It will be noticed that a reduction in the field causes a reduction in the peak values and a skewing of the position of the peaks in the direction of the unstable torque zeros; the positions of the zeros do not alter appreciably except at very small fields.

Closely associated with the error introduced when H is small is the error introduced by using a disk instead of an ellipsoid. In a disk the demagnetizing force is not uniform and very large fields are necessary to complete the alinement of I_s with H at the edges of the disk. If the slight skewing of the peaks which occur mainly below $H = 2000$ oersteds can be ignored, the effect of using a disk, and the effect of a relatively low field, can be allowed for by assuming lower values for the anisotropy constant.

Tarasov[7] showed that over the range of applied field H_a of 2000 to 3500 oersteds, the torque maximum depends linearly upon $1/H_a$, according to the relation

$$(T_H/T_\infty) = 1 - (H_0/H_a)$$

where H_0 is a constant. The value of H_0 depends on the orientation of the disk, the thickness/diameter ratio of the disk and the state of strain of the material. For polycrystalline cold-reduced 3% silicon-iron, Tarasov found that, averaged over the four torque peaks, the constant H_0 is given by

$$H_0 = -25 + 200 \log_{10}(1000\, t/D)$$

where $$3 < 1000\, t/D < 100$$

In Fig. 3 this data of Tarasov is plotted in the form of a family of curves, with K_{1a}/K_1 written as ordinate in place of

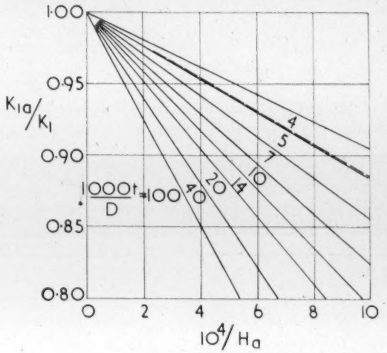


Fig. 3. Variation of the apparent value of the anisotropy constant with the applied field for annealed disks of polycrystalline 3% silicon-iron (Tarasov, 1939). The broken line applies to seven polycrystalline specimens of approximately (110) texture used in the present investigation

T_H/T_∞. K_{1a} can be regarded as the apparent value of the anisotropy constant to be used in place of K_1, to take into account the finite value of the field and the finite thickness of the disk.

Seven polycrystalline specimens of 35% silicon-iron, with appreciably different texture, as will be seen later, were tested in a similar way. The shape of the line was closely the same for each specimen, and is shown by the broken line of Fig. 3. The specimens each had a value of 1000 t/D of 14, and were all annealed prior to cutting the disk, and the disk was cut by a strain free method (electro-erosion). It will be noticed that the slope of this line is much less than would be expected from Tarasov's data; the reason for this is not known.

COLD-REDUCED SILICON-IRON SHEET

X-ray analysis has shown that cold-reduced silicon-iron sheet usually has a texture which is basically (110) with a spread about this occurring as a "roll" (α), pitch (β), and "yaw" (γ), as indicated in Fig. 4. There is sometimes also a proportion (Δ) of crystals with a random orientation. The spreads are approximately uniform over equal positive and negative limits.

The direction cosines defining the orientation of a single crystal at α, β, γ from the ideal orientation are

$$\left.\begin{aligned}\alpha_1 &= \sin\zeta\cos\phi - \cos\zeta\sin\beta\sin\phi,\\ \alpha_2 &= \sin\zeta\sin\phi + \cos\zeta\sin\beta\cos\phi, \text{ and}\\ \alpha_3 &= \cos\zeta\cos\beta,\end{aligned}\right\} \quad (1)$$

where $\zeta = \theta + \gamma$ and $\phi = 45° + \alpha$, the angles involved being defined in Fig. 4. If the spread about the mean

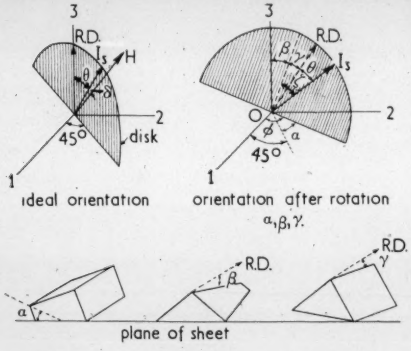


Fig. 4. Definition of angles employed in torque analysis of polycrystalline material with (110) mean orientation.
R.D. = rolling direction.

orientation is uniform about the three directions α, β, γ, over equal positive and negative limits, the expression for the mean torque is

$$T = -\frac{d}{d\theta}\frac{\int_{-\alpha}^{\alpha}\int_{-\beta}^{\beta}\int_{-\gamma}^{\gamma} E[(45°+\alpha), \beta, (\theta+\gamma)]d\alpha\, d\beta\, d\gamma}{2\alpha \,.\, 2\beta \,.\, 2\gamma} \quad (2)$$

$E[(45° + \alpha), \beta, (\theta + \gamma)]$ is obtained when α_1, α_2, α_3 in equation (1) are substituted in the well-known expression for the crystal anisotropy energy density

$$E = K_0 + K_1(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2(\alpha_1^2\alpha_2^2\alpha_3^2) \quad (3)$$

The expression for T on integrating and differentiating becomes

$$-T/K_1 = A_1\sin 2\theta + A_2\sin 4\theta + (K_2/K_1)(B_1\sin 2\theta + B_2\sin 4\theta - B_3\sin 6\theta) \quad (4)$$

where, if we write

$(\sin 2\alpha/2\alpha) = a_2$, $(\sin 4\alpha/4\alpha) = a_4$, . . . etc.,
$(\sin 2\beta/2\beta) = b_2$, . . . etc., $(\sin 2\gamma/2\gamma) = c_2$, . . . etc.,

then

$$A_1 = \frac{c_2}{2.32}[(-3 + 4b_2 + 7b_4) + a_4(5 + 4b_2 - b_4)],$$

$$A_2 = \frac{c_4}{4.32}[(21 + 28b_2 + 7b_4) + a_4(-35 + 28b_2 - b_4)],$$

$$B_1 = \frac{c_2}{4.32.32}[(-22 + 31b_2 + 38b_4 - 15b_6) + a_4(10 + 31b_2 + 6b_4 - 15b_6)],$$

$$B_2 = \frac{c_4}{32.32}[(18 + 19b_2 - 2b_4 - 3b_6) + a_4(-14 + 19b_2 + 30b_4 - 3b_6)],$$

$$B_3 = \frac{3c_6}{4.32.32}[(10 + 15b_2 + 6b_4 + b_6) + a_4(42 + 15b_2 - 26b_4 + b_6)].$$

The B terms are not very helpful in estimating the orientation, since K_2/K_1 is not known accurately and the B terms are small in comparison with the A terms. There is, therefore, no alternative but to assume K_2/K_1 to be zero and to deduce the orientation using the coefficients A_1 and A_2. With only two coefficients, the exact determination of $\alpha, \beta, \gamma, \Delta$ is out of the question unless additional relations between these quantities are known or assumed. Lack of knowledge of K_2/K_1 restricts the information which can be deduced from the torque curve, but, as in the case of orientation estimates with single crystals, it is not likely to be the cause of much error in that information.

As in the single crystal case considered by Tarasov and Bitter,[11] it is difficult to deduce the orientation given the coefficients, the more reasonable course being to calculate once and for all the coefficients for the whole range of orientations. The coefficients are not themselves particularly useful, since a harmonic analysis of the experimental curve is then necessary; it is better to express α, β, γ directly in terms of quantities easily measured from the torque curve. The two independent quantities, selected in place of A_1 A_2, might be the height of the major and minor peaks (T_1 and T_2, say) or T_1 and T_2/T_1, or T_1 and θ_{12}—the angle between torque zeros, or other combinations. There is little to choose between the various possibilities but T_1/K_1 and T_2/T_1 have been chosen and in Fig. 5 (*a* to *c*) the values of these quantities have been plotted over a wide range of α, β, γ. The abscissae in Fig. 5 are made $T_1/[K_{1a}(1 - \Delta)]$ rather than T_1/K_{1a} to allow for the simple inclusion of the case when random oriented material is present; and, as stated before, K_{1a} can be used rather than K_1 to allow for the demagnetizing effect in the disk and the finite value of the field.

Although the exact determination of α, β, γ is not possible from Fig. 5 unless a further relation is known, the different types of spread produce different effects on the shape of the torque curve which assist in their identification. It will be noticed from Fig. 5 that a spread occurring as a roll (α) tends to increase the minor peak without much change in the major peak; that is, it gives rise to a vertical displacement of the curve as a whole so altering T_2/T_1, but leaving the value of $T_1/[K_{1a}(1 - \Delta)]$ unchanged. Spread due to pitching (β) is evident as a reduction of the major peak but not so marked a reduction in the minor peak. Spread due to yawing (γ) is indicated by a reduction in the minor peak with a proportionally smaller reduction in the major peak. This is summarized in Table 1.

Table 1. *The effect of different types of spread on the shape of the torque curve*

Type of spread	*Major peak*	*Minor peak*
No spread (ideal orientation)	Normal ($T_1/K_{1a} = 0{\cdot}56$)	Normal ($T_2/K_{1a} = 0{\cdot}21$)
Uniform spread in all directions and/or proportion of random oriented material	Both peaks reduced in proportion	
Roll (α)	Slight reduction	Increased
Pitch (β)	Large reduction	Slight reduction
Yaw (γ)	Slight reduction	Large reduction

Since a spread occurring as a roll (α) is more of an advantage than a disadvantage, due to the improvement in the properties in a transverse direction without loss of properties in the rolling direction, then the change in the major peak is of greatest significance in the manufacture of cold-reduced silicon-iron sheet and indicates an undesirable spread about the ideal orientation.

The maximum quantitative information which can be extracted from a torque curve in the absence of information other than that the mean orientation is in the (110) plane, is given in Fig. 6. In this figure a new quantity is introduced, the "root mean square angle of spread," $\omega = (\alpha^2 + \beta^2 + \gamma^2)^{\frac{1}{2}}$. This angle gives a good idea of the deviation from the ideal in terms of a single quantity. Fig. 6 was deduced by plotting all the points at the intersection of the lines in Fig. 5 and drawing in the contours. The accuracy is not better than $\pm 10\%$ approximately, since a range of orientations, in which the component angles are different, but which are located by a single point in the figure, give rise to slightly different values of ω. Some knowledge of the relative values of α, β, γ is also possible using Fig. 6, as indicated by the marking of the different regions, and takes the information in Table 1 one stage further.

Incidentally, if the total angle ($\alpha + \beta + \gamma$) is used instead of the root mean square angle the accuracy is reduced to $\pm 50\%$, due to the total angle being over-emphasized when the individual angles are of the same order and being under-emphasized when one angle is large. The use of the root mean square angle overcomes this difficulty. A map similar to Fig. 6, but plotting $(\beta^2 + \gamma^2)^{\frac{1}{2}}$ instead of ω, should be of equal utility.

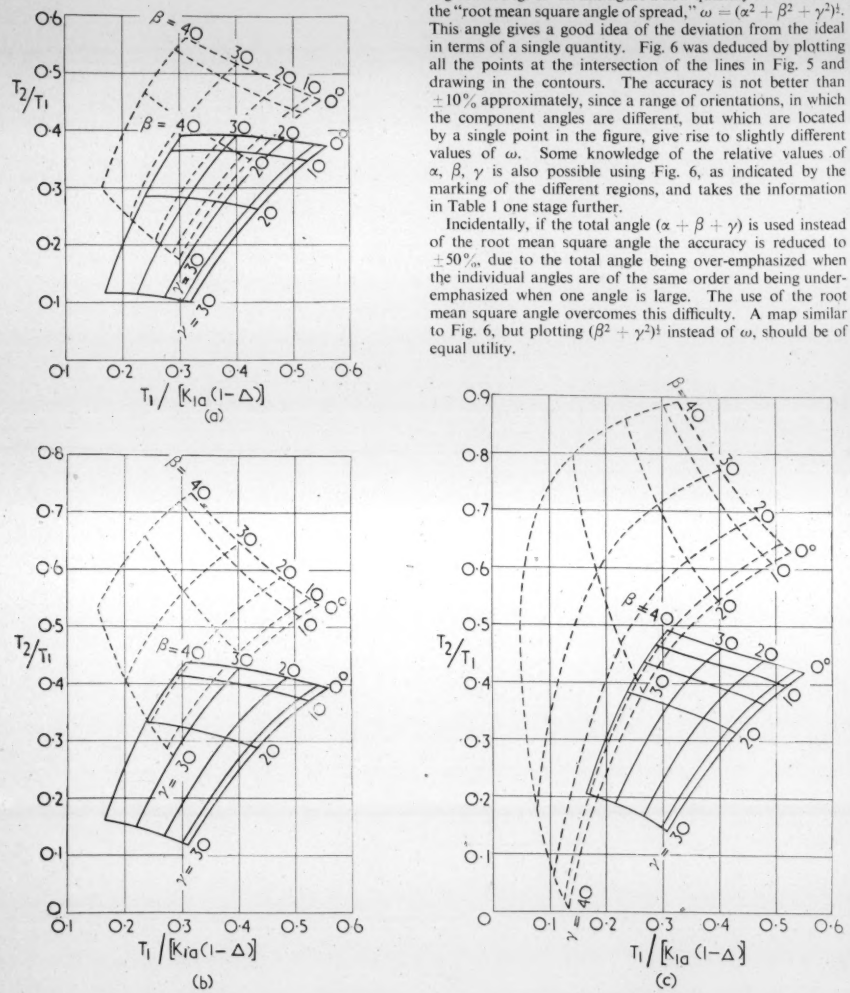


Fig. 5. Determination of the angles of spread (α, β, γ) and the proportion of random oriented material (Δ) from the magnitude of the torque peaks (T_1 and T_2). The curves are obtained assuming $K_2 = 0$

(*a*) Broken line curves $\alpha = 20°$ Full curves $\alpha = 0°$ (*b*) Broken line curves $\alpha = 30°$ Full curves $\alpha = 10°$ (*c*) Broken line curves $\alpha = 40°$ Full curves $\alpha = 15°$

Some idea of the errors introduced by neglecting the K_2 term is indicated by Fig. 7; the particular case is plotted of $\alpha = \beta = \gamma$, with K_2/K_1 zero and $+0{\cdot}35$. It is apparent that K_2 must be accurately known if the relative proportions of α, β, γ are to be deduced accurately when $\omega < 15°$. It rise to what is virtually three reflexions (a, b, c) from the three surfaces (A, B, C). The reflexions are plotted in a Mercator's projection [see Fig. 8(b)], modified in that the lines of longitude are a little distorted and the lines of latitude are closer together at the poles instead of further apart. With the polycrystalline material under discussion, the points a, b, c are spread over the areas shown shaded, the roll (α), pitch (β) and yaw (γ) being easily identified. Fig. 8(c) shows a texture map of a specimen of cold-reduced 3·5% silicon-iron with a well-defined texture. The specimen of Fig. 8(d) has a similar composition, but the texture is not so good and there is evidence of a random oriented component.

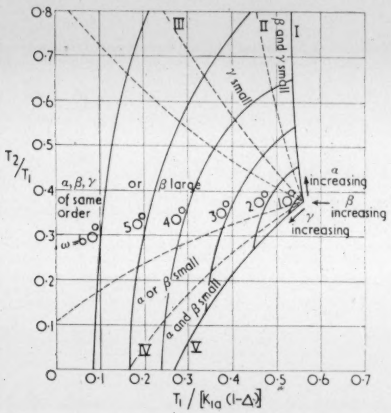


Fig. 6. Curve showing the variation of the "root mean square angle of spread" (U) with the torque maxima (T_1 and T_2). This figure gives the maximum information obtainable from a torque curve in the absence of any additional knowledge of the orientation

Curve I. $\beta, \gamma = 0$,
II. $\beta < \alpha/2$; $\gamma < \alpha/4$,
III. $\gamma < \alpha/2$,
IV. $\alpha < \gamma/2$; $\beta < \gamma/2$,
V. $\alpha, \beta = 0$.

will be seen also that the curves are parallel to the T_1/K_1 axis over most of the range, so that it is not possible to distinguish between the case of a small spread and a large random oriented component, and the case of a large spread and a small random component.

X-ray texture maps and torque curve data are compared in Table 2. The presentation of the information of the orientation in a texture map is a little different from usual, so two examples of texture maps are given in Fig. 8 (p. 98). As indicated in Fig. 8(a), a (110) [100] single crystal can give

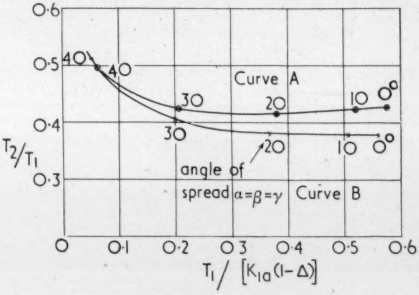


Fig. 7. Curve indicating the influence of K_2/K_1 for the case of equal angles of spread (i.e. $\alpha = \beta = \gamma$)

Curve A. $K_2/K_1 = 0{\cdot}35$.
Curve B. $K_2/K_1 = 0$

Details of these are included with the other specimens in Table 2.

In determining the orientation from the torque curves, K_{1a} was deduced from Figs. 1 and 3 using the broken curve in Fig. 3. T_1 and T_2 were measured from the torque curves. T_2/T_1 and T_1/K_{1a} were calculated, and the curves of Fig. 5 were used to obtain α, β, γ, Δ, those values being selected giving closest correspondence with the texture map data. Also tabulated is the value of ω from Fig. 6, together with the comments shown in the region in which the point T_2/T_1, T_1/K_{1a} is located.

In view of the errors which are known to be present—error in estimating from the texture maps, errors in K_{1a} and to assuming K_2 to be zero—there is remarkably good agree-

Table 2. *Comparison of X-ray texture map and torque curve data*

Specimen identification	*X-ray texture map data*					*Torque curve data*		*Deductions from torque curve*				
	% Random Δ	*Roll* α	*Pitch* β	*Yaw* γ	$\omega = (\alpha^2 + \beta^2 + \gamma^2)^{\frac{1}{2}}$	T_2/T_1	$T_1/3{\cdot}3 \times 10^5$	*Using Fig. 5* Δ	α	β	γ	*Using Fig. 6* ω
26	small	10	12	10	18	0·39	0·51	0	10	14	8	18 (same order)
27	?	15	15	15	26	0·36	0·48	0	10	12	14	20 (same order)
56	small	14	13	16	25	0·40	0·49	0	15	12	12	21 (same order)
55(a)	small	10	10	22	26	0·26	$0{\cdot}43_5$	0	0	0	23	22 (α and β small)
55(b)	small	12	12	20	27	0·29	$0{\cdot}47_5$	0	0	0	20	18 (α and β small)
28	*	15	large	15	large	0·40	0·41	0	15	23	14	32 (same order)
57	*	15	large	20	large	0·31	0·25	0	15	34	24	45 (same order)
								26	15	20	22	

* May be interpreted as a large random component or a large spread occurring as a pitch (β).

Test details. Specimen 0·95 in. diameter, 0·014 in. thick, 3·4% Si or near. Field 1930 oersteds.
From Fig. 1, $K_1 = 3{\cdot}5 \times 10^5$. From Fig. 3 (broken line) $K_{1a} = 3{\cdot}5 \times 10^5 \times 0{\cdot}94 = 3{\cdot}3 \times 10^5$.

ment between the X-ray texture map and torque curve data. This correspondence was only achieved, however, when the disk was cut out of the strip by a strain-free method (electro-erosion), keeping the variations in the diameter of the specimen to $\pm 0{\cdot}001$ in. The importance of errors due to using non-circular disks was pointed out in a private communication from D. A. Langford, who showed that a 1% difference between perpendicular diameters could alter the maximum torque T_1 by as much as 17% and the ratio T_2/T_1 by 25%. Strain-free cutting is important in that it controls the slope of the K_{1a}/K_1 curve of Fig. 3, and so gives rise to errors in estimating K_{1a}.

CONCLUSIONS

It is not possible in general to deduce from a torque curve an accurate picture of the texture of polycrystalline material, but, if the texture can be summarized in a simple manner—preferably by two parameters—then the torque curve method is simple and accurate. Cold-reduced silicon-iron usually has a mean orientation of (110) [100] with spreads approximately uniform over the angle $\pm \alpha$, $\pm \beta$ and $\pm \gamma$ in mutually perpendicular directions respectively about the mean; there is also sometimes a random oriented component Δ. It is not possible to deduce α, β, γ and Δ from a torque curve unless two additional items of information are known. If this is provided by X-ray methods, then the torque curve enables the X-ray data to be simply extended and expressed more accurately. Interpretation of the torque curve is rapid and requires only the measurement of the major and minor torque peaks when the texture can be obtained from charts. If nothing is known other than that the mean orientation is (110) [100], the order of magnitude of the spreads can be deduced in terms of a single quantity $\omega = (\alpha^2 + \beta^2 + \gamma^2)^{\frac{1}{2}}$, and some indication of the component angles is given vaguely in general terms, such as α large, or α and β small, etc.

ACKNOWLEDGEMENTS

I would like to acknowledge the assistance of Miss M. Marshall in the computations, Mr. B. K. Ginty for torque magnetometer measurements, and Mr. J. A. James for the X-ray analyses. I am also grateful to my colleagues for valuable discussions, and to Mr. L. J. Davies, Director of Research, the British Thomson-Houston Company Ltd., for permission to publish this paper.

REFERENCES

(1) BITTER, F. *Introduction to Ferromagnetism*, pp. 213–22 (New York: McGraw-Hill Book Co. Inc., 1937).
(2) MILNER, C. J., and JAMES, J. A. *J. Sci. Instrum.*, **30**, p. 77 (1953).
(3) AKULOV, N., and BRUCHATOV, N. *Ann. Phys.*, **15**, p. 741 (1952).
(4) BRAILSFORD, F. *J. Instn Elect. Engrs*, **83**, p. 566 (1938).
(5) DE BARR, A. E. *J. Instn Elect. Engrs*, **96**, Pt. II, p. 719 (1949).
(6) BOZORTH, R. M. *Phys. Rev.*, **50**, p. 1076 (1936).
(7) TARASOV, L. P. *Phys. Rev.*, **56**, p. 1224 (1939).
(8) BOZORTH, R. M. *Ferromagnetism*, p. 572 (New York: McGraw-Hill Book Co. Inc., 1951).
(9) WILLIAMS, H. J. *Phys. Rev.*, **52**, p. 747 (1937).
(10) TARASOV, L. P. *Phys. Rev.*, **56**, p. 1231 (1939).
(11) TARASOV, L. P., and BITTER, F. *Phys. Rev.*, **52**, p. 353 (1937).

Some observations on the time edge effect in Catalin 800

By A. R. MORRIS, M.Sc., National Physical Laboratory, South African Council for Scientific and Industrial Research, Pretoria

[*Paper first received* 12 *January, and in final form* 10 *August*, 1953]

The investigation shows that the "time edge" effect in the phenol-formaldehyde type of resin is accelerated by exposure to oxygen and moisture and retarded by storage at low temperatures. It is also noted that the time edge effect apparently proceeds in two stages.

The time edge effect is always a source of inaccuracy in photoelastic analysis, but the textbooks have very little information as to its avoidance beyond advising the reader to use a photoelastic model within a few hours of the final trimming of its edges. Some success has been found in delaying the time edge effect in the glyptal type of resin (Bakelite BT 61–893) by keeping it in a desiccator, but this method is not so successful in preserving the edges of models made from the phenolic resins (Catalin 800).

Because of the scarcity of published information on the subject, an investigation was carried out in this laboratory with a view to finding some of the factors which accelerate or retard the development of time edge stresses in Catalin 800, and thus determining the best way to use this material. Except in the case where storage in a refrigerator is mentioned, all the experiments were done at room temperatures. Thus the observations are rather different from those of Stevens and Marshall,* who have examined the time edge effect in Catalin heated in various liquids to above 80° C as is necessary for the frozen stress method.

The fringe order at the edge of the material was chosen as the most convenient index of the progress of the time edge effect as it may be regarded as proportional to the tangental stress produced by the effect. The chief difficulty introduced into photoelasticity by time edge effect is this small stress which makes the accurate determination of the edge stress of the pattern being studied difficult and uncertain.

One series of experiments consisted of exposing a number of freshly cut Catalin blocks to a variety of different conditions and photographing them at intervals. The photographs were enlarged and the number of fringes was counted at the edge to measure the progress of the time edge effect. For convenience four blocks were dealt with at a time and for comparison they were piled on top of one another and photographed simultaneously. Nine photographs were made on each 4×5 in. negative, using a set of masking cards with a $\frac{1}{2}$ in. wide vertical slot. Fig. 4 (p. 98) shows a representative photograph.

Some typical results are shown in the accompanying graphs, in which the fringe order at the edge is plotted against time

* STEVENS, P. R., and MARSHALL, T. A. *Distrib. Elect.*, October 1948 and January 1949 (London: W. T. Henley's Telegraph Works Co. Ltd.).

in days. As can be seen, the development of time edge effect is accelerated by the presence of either moisture (Fig. 1, curve *C*) or oxygen (Fig. 3, curve *E*). The presence of both water and oxygen together as in curve *G*, Fig. 3, when the

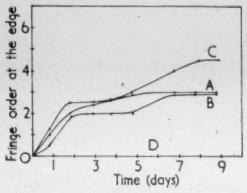


Fig. 1. Fringe order at edge *versus* time in days

A, in air; *B*, in a dry atmosphere; *C*, in a wet atmosphere; *D*, in a refrigerator.

Catalin was immersed in hydrogen peroxide, accelerated the time edge effect so much that the fringes were difficult to count after the first two days.

The edge stress in sample *B*, Fig. 1, which was kept in a desiccator containing silica gel was slightly retarded as compared with sample *A*, Fig. 1, which was kept in air, but it was not enough to justify the general use of the desiccator as a means of preserving Catalin models. The piece of Catalin,

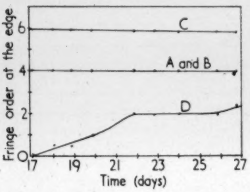


Fig. 2. The same blocks as in Fig. 1, but removed from their special environments one week later and exposed to the air

D, Fig. 1, which was kept in a refrigerator (about $-8°$ C) showed no time fringes at all even though no attempt had been made to protect it from air or moisture.

A week after the experiment, illustrated by Fig. 1, had been stopped, the pieces of Catalin were removed from their special environments and exposed to the air for the further series of photographs, the results of which are illustrated by Fig. 2. The block *D* which had been in the refrigerator,

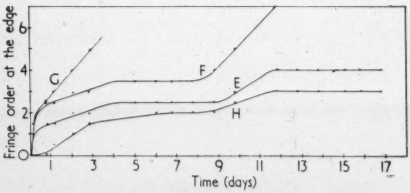


Fig. 3. Fringe order at edge *versus* time in days

E, in dry oxygen; *F*, in water; *G*, in hydrogen peroxide; *H*, in air.

slowly developed edge stresses, taking five days to reach the same condition that was reached in two days by the block *A* kept under normal conditions (about 25° C). Thus the effect of the low temperature is not only to retard time edge effects while the Catalin is in the refrigerator, but also to slow down the subsequent development of edge stresses after removal from the refrigerator. Provided the model is used within about a week after trimming the edges, time edge effects are not noticeable.

In the many experiments performed on the time edge effect, very nearly all the graphs show a characteristic upward bend after four or five days. Following the initial rapid rise the curves become almost horizontal, with a further slow rise after a few days. This suggests that the time edge effect in this kind of plastic proceeds in two stages. The first is probably an oxidation or moisture absorption process and the second possibly a change in the state of polymerization of the material caused by the first stage. The first sudden rise is usually steepest in those specimens subjected to moist conditions and these also develop the characteristic green colour of old Catalin.

The experiments illustrated by the above-mentioned graphs were performed in a room which was not air conditioned or temperature-controlled and the temperature was therefore a variable factor. The results were not perfectly repeatable. An experiment done in the summer and later repeated during the winter would produce graphs of similar form but of lower values of fringe order at each stage. Another variable which was found to be a factor was the age of the Catalin. Fresh supplies were found to develop edge effects more rapidly than old matured stocks.

A large number of experiments have been carried out to determine the effects of refrigeration on the photoelastic sensitivity of the material. Provided the blocks are allowed to warm up to room temperature before doing a test, it appears from the results that the material fringe value of Catalin is unchanged by storage at a low temperature.

It is not claimed that storage in a refrigerator is the perfect solution to the problem of avoiding time edge effects. Pieces of Catalin kept in the refrigerator for more than a month show fine edge fringes. Such models may be good enough for demonstration purposes, as the fringes at the edge are difficult to notice, but they are not good enough for accurate photoelastic work.

A piece of Catalin which has been stored for some time in a refrigerator shows a uniform bright colour, when viewed in a crossed circular polariscope, instead of being uniformly dark as is usual in the completely unloaded condition. This effect gives rise to the fear that the zero orders of a photoelastic pattern might be lost or displaced if the model has been stored in a refrigerator. To put this possibility to the test, a well-known photoelastic pattern, which has a well-defined zero order fringe (a bent rectangular beam), was set up and photographed. The model was then stored in a refrigerator for three months and then again photographed under identical conditions of load and magnification. The zero order fringe (neutral section) appeared to be still in its proper place, but on superimposing the two negatives it could be seen that in the second case the zero order fringe was very slightly displaced towards the compression side of the beam. The displacement was of the order of one-fifth of the spacing between fringes. It would appear that refrigeration must be used with caution with models whose patterns contain important zero order features such as isotropic or singular points.

A consideration of the effects of dry and wet atmospheres raises the question of the time edge effect being a process in which water is lost or a process in which water is absorbed. To put this to the test a number of small pieces of Catalin, approximately $\frac{1}{4} \times 1 \times \frac{1}{16}$ in., of which most of the surface area could be regarded as freshly cut edge, were weighed and then subjected to the conditions used in previous experiments. After a week they were weighed again and both gains

and losses of weight were noticed. The pieces which had been kept in the damp atmosphere had gained weight equal to slightly more than 3% of their total weights. The pieces which had been kept in a desiccator containing silica gel had lost slightly more than 3% of their weight. The pieces which had been exposed to the air had lost rather less than 1% of their weight. The relative humidity in the room was, of course, a variable factor, but most of the time during this experiment it was between 40 and 65%. The small pieces which had been kept in the refrigerator, although covered with ice, had gained less than 1% of their weight.

The width of the band of fringes is sometimes used as a rough guide to the progress of the time edge effect. This apparent depth of penetration of the effect was investigated by measuring the fringes in the photographs with a travelling microscope. The results were similar to those of the fringe count at the edge as regards the effectiveness of retarding or accelerating factors, but they did not produce graphs of the shape shown in Figs. 1 and 3. The depth of penetration appears to increase approximately linearly with time.

The measurement of fringe band widths gives very much more irregular results than fringe counts, one reason for this being the fact that there is often a discrepancy between the depth of penetration as defined by the distance from the edge to the zero order, and the width of the narrow band of fringes at the edge. The zero order usually shrinks to occupy a narrow belt along the centre of the block and after a few weeks it may be only represented by two small black dots. In the experiments previously described the width of the fringe band after a week is usually of the order of $\frac{1}{16}$ in. or less. Examination of large sheets of Catalin several months old may show edge effects occupying a band more than an inch wide although there may be still only three fringes. The small $1 \times \frac{1}{4}$ in. blocks are clearly not suitable for investigating the depth of penetration of the time edge effect thoroughly.

ACKNOWLEDGEMENTS

The experimental work described in this paper has been carried out in the National Physical Laboratory, Pretoria, and the author wishes to thank the South African Council for Scientific and Industrial Research for permission to publish the paper.

The correction of an error in the determination of high resistances by capacitor discharge

By G. C. Curtis, M.A., Ph.D., A.Inst.P.,* University College, London

[*Paper first received* 28 *September and in final form* 19 *October*, 1953]

In the course of determinations of the surface resistivity of hard-glass and silica rods from observations of the discharge of a Wulf electrometer, it was found that the voltage decay was initially non-exponential. This is attributed to the effect of capacitance between rod and case, distributed along the length of the rod. The effect is treated mathematically and it is shown that, even if the resistance is determined from the exponential portion of the curve, a correction is required. A formula for the correction is given and an approximate method of applying it referred to. An indication is given of the circumstances, probably not uncommon, in which values of resistance determined by the leakage method, might require appreciable correction.

High resistances are commonly determined by using them to discharge a capacitor of known value, observing its potential as a function of time. In the experiments described, a marked departure from the exponential law was found. This is quantitatively accounted for as due to capacitance to earth being distributed along the resistor. If this capacitance is comparable with the "known" capacitance, the theoretical treatment given shows that a considerable correction to the calculated value of the resistance is required.

EXPERIMENTAL DETAILS

Although certainly of wider occurrence, the effects to be described were actually observed during determinations of the surface resistivity of glass rods. The surface resistivity was being studied as a function of humidity, but in each experiment cited this parameter was held constant. The hygrometer and thermometer are omitted from the simplified sketch (Fig. 1). A brass chamber encloses the rod under test, which forms the sole support of a Wulf electrometer. This electrometer consists essentially of two fine metallized quartz fibres held in tension by a quartz bow; the separation of the fibres is observed through a microscope fitted with a graticule. After calibration, this separation is a fairly accurate measure of the potential V of the lower end of the glass rod, to which the fibres are electrically connected. The electrometer is charged by momentarily connecting it to a battery through a pivoted "hairpin" of iron wire, magnetically operated from outside the case. The iron wire is then earthed so that its effect on subsequent readings is always the same.

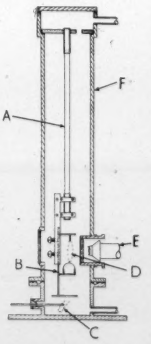


Fig. 1. Apparatus for determination of resistance by leakage

A, glass rod; *B*, Wulf electrometer; *C*, "hairpin" of iron wire; *D*, metallized quartz fibres; *E*, microscope; *F*, brass chamber.

Conduction through the air and the solid glass can be shown to be negligible; hence, the electrometer can only lose its charge by conduction over the surface of the glass rod to the earthed clamps at the top.

* Now at Department of Atomic Energy, Sellafield, Cumberland.

THEORY

Deviations from simple decay law. In many cases, the linearity of the plot of log V against time (see AB in Fig. 2) showed that Ohm's law was obeyed; from the slope, the resistance R is calculable, since

$$V = V_0 \exp(-t/CR) \tag{1}$$

where the electrometer potential is V_0 at time 0, and V at time t. The capacitance C between electrometer and case was determined as $7{\cdot}6 \times 10^{-12}$ F by both a.c. and d.c. methods.

If V is found to be halved in T sec then from equation (1):

$$R = (1/C \ln 2)T. \tag{2}$$

On the other hand, particularly when T was large, initial departure from linearity of the plot was sometimes noticed (see CD in Fig. 2). This was traced to the fact that the

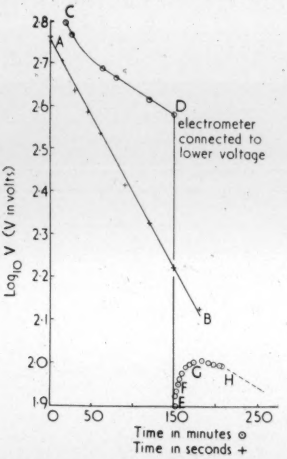


Fig. 2. Logarithmic voltage curves. The point D indicates the time when the electrometer was connected to a lower voltage

surface of the rod has a capacitance K per unit length. Let its resistance per unit length be $r = R/L$, where L is the length of rod between the contacts. Fig. 3 shows the equivalent circuit.

If, at a point on the rod distant x from the electrometer, the potential is v at time t, elementary transmission line theory [see ref. (1)] gives

$$\frac{\partial v}{\partial t} = \frac{1}{Kr}\frac{\partial^2 v}{\partial x^2} \tag{3}$$

Boundary conditions are required. The line is terminated at $x = 0$ by a capacitance C; equating expressions for the current at this point, $\frac{1}{r}\frac{\partial v}{\partial x} = C\frac{\partial v}{\partial t}$; the line is earthed at $x = L$, where $v = 0$ always. When $t = 0$ (the instant of charging) $v = V_0$ at $x = 0$ and zero elsewhere.

The appropriate solution, given (in connexion with a heat problem, and with different notation) by Carslaw and Jaeger[2] is

$$v = V_0 \sum_{n=1}^{\infty} \frac{2U}{\alpha_n^2 + U^2 + U} \frac{\sin \alpha_n\left(1 - \frac{x}{L}\right)}{\sin \alpha_n} \exp\left(-\frac{\alpha_n^2}{KrL^2}t\right) \tag{4}$$

where $U = KL/C$ and $\alpha_1 \alpha_2 \ldots \alpha_n \ldots$ are the roots of $\alpha \tan \alpha = U$.

At the electrometer $x = 0$ and $v = V$ by definition.

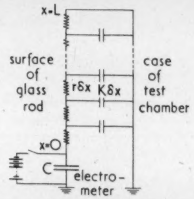


Fig. 3. Equivalent circuit for conduction over surface of rod

Example. Values of the quantities appropriate to the apparatus described will now be introduced. For this, $C = 7{\cdot}6 \times 10^{-12}$ F (see above) and $L = 18$ cm. By regarding rod and case as a concentric-cylinder condenser, K was calculated as $0{\cdot}25 \times 10^{-12}$ F.cm^{-1}. Equation (4) then reduces to

$$V = V_0[0{\cdot}827 \exp(-0{\cdot}693\, t/T) + 0{\cdot}098 \exp(-15{\cdot}7\, t/T) + 0{\cdot}028 \exp(-58\, t/T) + \ldots] \tag{5}$$

where T is defined as $KrL^2(\ln 2/\alpha_1^2)$.

Therefore the slope of the curve of log V against t should decrease rapidly for $0 < t < 0{\cdot}1\, T$, but afterwards approach a constant value such that V is halved during any interval T (since $0{\cdot}693 = \ln 2$) in agreement with the earlier definition of T. The curve CD of Fig. 2, calculated theoretically, predicts values of V disagreeing by at most 0·6% from experiment: this was achieved by using only one parameter, T, derived from the curve itself (the straight portion). T merely determines the scale of the t-axis. (The starting-point is, of course, experimental.)

In determining R from the asymptotic slope we have

$$R = rL = (\alpha_1^2/KL \ln 2)T \tag{6}$$

Thus the value of R obtained from equation (2) should be multiplied by $C\alpha_1^2/KL$ which is 0·832 for this particular geometry of apparatus. Its neglect by earlier workers with this type of apparatus may have introduced considerable errors in cases where the ratio of distributed to lumped capacitance was appreciable. This is especially likely to happen if high current sensitivity is being sought by using an electrometer of small capacitance. It is true that attention should be directed to this error by curvature of the logarithmic plot, but if the decay time were, say, 2 min or less this might well be overlooked.

Where only an approximate correction for the distributed capacitance is required, it is sometimes permissible to regard it as equivalent to a lumped capacitance of one-third of its total value. In the present instance this is one-third of $18 \times 0{\cdot}25 \times 10^{-12}$ F and is to be taken as being in parallel with C, $7{\cdot}6 \times 10^{-12}$ F, giving a total of $9{\cdot}1 \times 10^{-12}$ F.

The correction factor for R is thus 7·6/9·1 or 0·835, differing by only 0·4% from the value given above. This approximation is, however, only valid if the voltage distribution remains uniform during the period of time considered.

Finally, confirmation of the hypothesis that the rod can store charge was obtained by allowing the electrometer to leak till linearity was established (see point D in Fig. 2) and then momentarily connecting it to a source of lower potential (point E). After re-isolation, charge leaked back into the electrometer from the rod (see FG in Fig. 2) causing its potential to rise (as predicted) before normal leakage supervened (see H in Fig. 2).

ACKNOWLEDGEMENT

I wish to thank Prof. E. N. da C. Andrade for providing the necessary facilities at University College, London, where the work was carried out.

REFERENCES

(1) JEANS. *The Mathematical Theory of Electricity and Magnetism*, pp. 332–3 (London: Cambridge University Press, 1933).

(2) CARSLAW and JAEGER. *Conduction of Heat in Solids*, p. 106 (London: Oxford University Press, 1947).

Some effects of slow electron bombardment in thermionic valves*†

By D. A. WRIGHT, M.Sc., F.Inst.P., M.-O. Valve Co. Ltd., at the G.E.C. Research Laboratories, Wembley, Middlesex

[*Paper first received* 14 *July, and in final form* 25 *September*, 1953]

This paper describes first the general features of the interaction between slow electrons and solids upon which they impinge. The threshold electron energies for the absorption of electrons, for the appearance of secondary emission and for decomposing the bombarded solid have been measured for a number of compounds. The values are quoted and are related to the optical absorption bands of the compounds. The effects on valve characteristics and on cathode emission are described briefly, when such compounds are present on the valve electrode surfaces.

When slow electrons bombard solid surfaces, such as those of valve electrodes, the following possibilities exist:

(1) Some of the electrons will be reflected without loss of energy (elastically).

(2) Some will collide with or enter the solid, and be reflected inelastically, having given up some energy to the solid.

(3) Of those which lose energy to the solid, some may be captured at its surface, if an insulating film is present. If there are many of these, then subsequently the number in the first group will increase because of the negative charge.

(4) Some will enter the solid, and continue moving in it until their energy is lost.

The first and second groups will obviously be of interest in a valve. If another electrode is present at a potential higher than that of the bombarded electrode, the effect of reflexion on current distribution will obviously be very considerable. Even in a diode, however, electrons reflected from the anode may affect the characteristics through the space charge they create. The time of travel of the reflected electrons to wherever they are collected may be important, as may be the noise they produce. The third group will be effective in altering the potential of the electrode, again modifying the characteristics. There are several consequences of the transfer of energy from the bombarding electrons to the solid:

(1) They will share some of their energy with the lattice vibrations, raising the temperature of the solid.

(2) They may induce chemical decomposition if the electrode material is not an element. If it is an element, they may yet produce structural changes by moving atoms from their normal lattice positions.

(3) They may produce excitation or ionization of the ions of the solid. Thus electrons may be freed from these ions: then (*a*) if the energy of these electrons is sufficient for them to reach the surface of the solid and to escape from it, they are emitted as secondary electrons. If these emerge from the bombarded surface they join the reflected electrons and influence the characteristics in a similar way to them, though their velocity distribution will be quite different. These considerations apply both to metals and non-metals. If the solid is not metallic, i.e. has low electrical conductivity before bombardment, then (*b*) an increase in conductivity may be detected on bombardment due to the freed electrons which do not escape—this is bombardment induced conductivity, analogous with photoconductivity. It will not be of great interest in valves, except in its implication that the resistance of thin, poorly conducting films under bombardment is rather indeterminate. The third possibility (*c*) also concerns electrons freed in the solid, but not lost from it. Ultimately these are recaptured. If suitable "centres" are present in non-metallic solids, luminescence may occur on recapture, as in cathode-ray tube screens. Luminescence is frequently detected on electrodes contaminated with cathode deposits, and sometimes on the glass envelope. It is of little importance in ordinary valves. The fourth possibility (*d*) is that the low energy bombardment will produce soft X-rays, which can be of some importance in valves since they will liberate electrons from any other solid which they strike. This is noticed primarily in ionization gauges, where X-rays from the anode liberate electrons from the grid, so setting a limit below which the positive ion current to the grid cannot be measured.(1)

The decomposition of compounds is of considerable importance in valves, in two respects. Firstly, all useful secondary emitting target materials are compounds, and their decomposition under bombardment leads to an instability in their secondary emission coefficient. This has been the main difficulty in the application of secondary emission to thermionic devices (it can be employed readily in photoelectric devices because of the lower current density, at which the instability is less marked). The decomposition is important secondly because it frequently leads to the liberation of the gaseous constituent. In extreme cases this can cause softening of the valve, and more frequently it causes "poisoning" of the emission without a large increase in pressure. It is important to understand that decomposition of this type can occur without a large temperature rise in the bombarded material. If there is a temperature rise then, of course, there will also be a degree of decomposition depending on the dissociation pressure of the compound at the tem-

* Communication No. 561.

† Based on a paper read on 29 May, 1953, before the Electronics Group at The Institute of Physics Convention in Bournemouth.

perature reached, but even with no temperature rise there can be a direct transfer of energy from the bombarding electrons to the crystal lattice, causing electron transitions and so a destruction of the ionic part of the chemical bond. Following this the gaseous atom may be able to escape. This is why electrodes can be well outgassed by heat treatment and yet liberate gas at a lower temperature on subsequent electron bombardment. Clearly there will be a minimum electron energy necessary to promote the effective transitions.

ELASTIC REFLEXION

In this and the following sections some of the features outlined above will be discussed in greater detail, and illustrations will be given based largely on recent work in this laboratory. Reference is also made to earlier relevant work.

There have not been many careful measurements of the reflexion of slow electrons. A recent study of electrons of thermal energy was made by Myers[2] working on copper, silver and tungsten. He found reflexion coefficients near 10% at energies below 5 V, i.e. 10% of all the incident electrons were reflected if their energy was 5 V or less. This proportion is of the same order as the theoretical figure for clean metals. When layers of poorly conducting material cover the metal two things happen: (*a*) the reflexion coefficient is usually greater than from a clean metal and (*b*) some electrons are trapped at or near the insulator surface. The latter may, of course, be the cause of the former. In recent work here[9] we have found that in a simple diode with a thin film of barium oxide (about 10^{-5} cm) on the anode, when 2 or 3 V are applied to the anode, the current decays below its initial value. This is due not to a fall in emission, but to a decrease in the potential at the outer surface of the film, by a few tenths of a volt. Typically, about 10^{10} electrons/cm^2 become trapped at the film surface under these conditions.

In the course of this work, we have made rough estimations of the reflexion coefficient at a barium oxide surface and find a value about 9 times that of nickel for electrons of 1 eV energy, and 3 times that of nickel at 3 eV, i.e. about 90% and 30% respectively. Rather high values are perhaps to be expected in view of the low value of A in the Richardson equation for barium oxide and similar compounds.

ABSORPTION

If the energy of the bombarding electrons is raised from zero, a stage is reached at which the reflexion decreases abruptly. This is the threshold for inelastic collisions, i.e. for energy transfer to the solid. The threshold can be detected in the diode characteristic, again for an anode with a thin film of barium oxide of thickness 10^{-5} cm. This is illustrated in Fig. 1. Here curve A shows a diode characteristic for a clean nickel anode, and is of the normal type. When barium oxide is present on the anode, curve B, the retarding field section of the curve is displaced to the left, because of the change in contact potential. Moreover, between $V_A = 2$ and $V_A = 4$ volts the curve B is lower than curve A, because of increased reflexion at the anode. A kink is detected at $V_A = 5$ volts, where absorption of electrons in the barium oxide commences. This value of V_A is not corrected for contact potential, and represents an electron energy of 4 eV.

This absorption lowers the space charge density and raises the anode current. The effect is perhaps better shown in a triode operated with the grid positive, and the anode voltage increasing from zero. Fig. 2 shows the grid current when the grid is held at +75 volts, the cathode is at earth (zero) potential, and the anode potential rises from 0 to 16 volts. The broken curve illustrates the usual behaviour when anode and grid are metallic. The grid current at first falls as anode current increases, until near 7 volts secondary emission from the anode reverses the slope. With a film of barium oxide on the anode the trend is the same, but kinks are detectable. The current at first falls, though more slowly than with a metallic anode since the reflexion is greater. Then at 4 volts the absorption occurs. On increasing the anode voltage, further changes in slope occur in the grid current plot, e.g. the increases at energies of 5, 7 and 10 eV. Further kinks have been observed at higher energies, but we have not been able to study them systematically, as decomposition effects produce

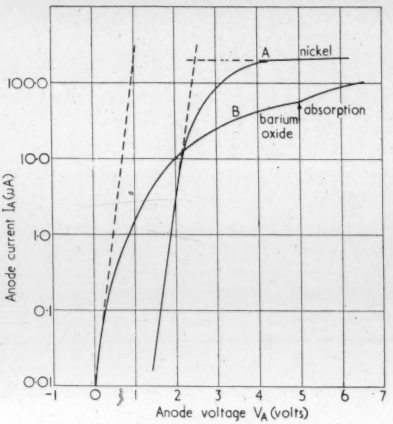


Fig. 1. Retarding field characteristics for a diode
Curve A: a clean nickel anode; Curve B: a nickel anode covered with a film of barium oxide.

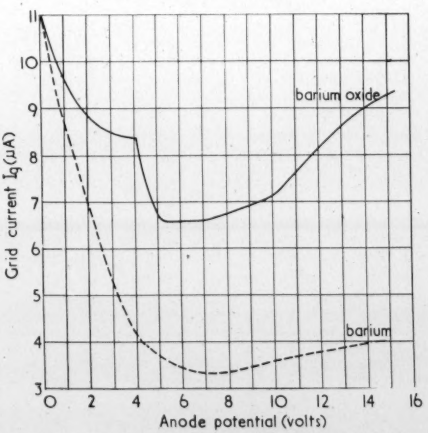


Fig. 2. Grid current *versus* anode potential of a triode with grid at +75 volts. The anode potential is measured relative to the cathode and corrected for contact potential

a marked instability. Such kinks are to be expected, however, as the electrons suffering inelastic collisions with the film on the anode can only dissipate definite quantities of energy because of quantum processes. This is well illustrated by the work of Rudberg[3] who studied the distribution in energy of electrons of a few hundred volts energy reflected from various surfaces. He found, for barium oxide surfaces, peaks in the distribution of reflected electrons at energy losses of 7 and 10 eV.

SECONDARY EMISSION THRESHOLD

The first increase in slope of the grid current in Fig. 2 occurs at the primary energy at which secondary emission becomes detectable, causing an apparent increase in reflexion.

Although the slope does not become definitely positive, the first increase in slope occurs at 5 volts, shortly after the absorption sets in at 4 volts. It seems likely that secondary emission is actually initiated here, and intensified at 7, 10 and 14 volts. The work of Apker and others[4] has shown that with barium oxide the filled band is at a depth of 5 eV below the energy of an electron outside the solid, so that this energy should allow some secondary emission. It is not yet clear why the marked increase should occur at 7 volts. There is little doubt that here a further absorption effect occurs, although it is usually masked in plots such as Fig. 2 by the increase in secondary emission; 7 eV is in fact the first energy loss observed by Rudberg. The secondary emission coefficient, as is well known, rises as the energy continues to increase, reaching a maximum of 6–7 at 5–600 volts for barium oxide.

We have measured the secondary emission threshold for various solids, and some values are listed in Table 1. Some metals are included here. If any of these materials are present on valve electrodes, corresponding kinks in the characteristic may be observed.

Table 1. *Secondary emission thresholds*

	eV				*eV*	
Ba	8	W		RbCl	9·1	Ja
BaO	7	W		CsCl	10	Ja
BaS	9	W		Cs_2O	8	Ja
$BaCl_2$	10·5	W		Cs_2O–Ag	7	Ja
$BaSO_4$	12	W		MgO	10	L
SrS	13	W				
Ni	8·5	W				
NaCl	10	W	8 Jo	10·5	H	
KCl	10·7	Ja	10	H		
KI	8	H	6·5	Jo		

H refers to Hilsch, ref. (6).
Ja refers to Jacobs, ref. (7).
Jo refers to Jonker, ref. (13).
L refers to Lempicki, ref. (12).
W refers to Wright and Woods, ref. (9).

DECOMPOSITION

Reverting to the first absorption threshold, this occurs when the bombarding electrons attain an energy of 4 eV, as shown in Figs. 1 and 2. It is noteworthy that optical experiments on either thin films or single crystals of barium oxide show that while the transparency is higher at longer wavelengths, absorption commences and rises rapidly as the wavelength of incident radiation falls below 3250 Å; thus there is a threshold here, corresponding with an energy of 3·8 eV.[5] Hilsch observed similar effects with some of the alkali halides,[6] while recent experiments in this laboratory confirm and extend his results. Thus if a triode is used with the anode contaminated with one of the alkali halides, the grid current shows the same type of behaviour as in Fig. 2, but the thresholds are of course different. Electron absorption becomes large in sodium chloride at 7·7 eV, in potassium chloride at 7·0 eV, and in potassium iodide at 5·5 eV. These energies are also similar to those at the respective optical absorption thresholds. Fig. 3 shows for these halides and for barium oxide curves giving the spectral distribution

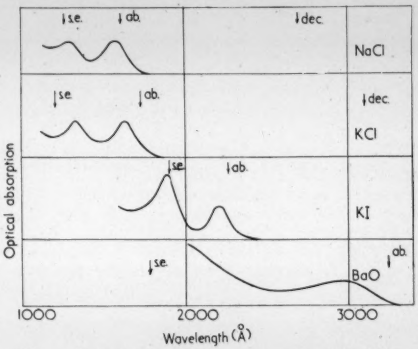


Fig. 3. Optical absorption *versus* wavelength for some halides and barium oxide. The decomposition of potassium iodide sets in at 2·5 eV, equivalent to 4900 Å

of absorption of incident radiation, while the arrows marked "ab." show the energy thresholds for electron absorption determined from the triode experiments, plotted with an energy scale corresponding with the wavelength scale. The arrows marked "s.e." show the energies where the slope of the grid current becomes abruptly positive. These are the energies where the secondary emission rises rapidly. They appear with the halides to be related to a second optical absorption maximum. This is to be expected if the second maximum corresponds with the freeing of an electron into the conduction band, as is often supposed, whereas the first maximum corresponds with exciton formation. With barium oxide the optical measurements have not been taken to wavelengths short enough to show whether a second maximum is present.

The arrow marked "dec." for each compound refers to the threshold for decomposition of the film on the anode under electron bombardment. This is determined in experiments with diodes with the cathode at a low temperature, e.g. 400° C. Work of this type was done by Jacobs,[7] and we have recently studied some of the films in great detail. With barium oxide, of thickness 10^{-6}–10^{-5} cm, instability of emission sets in when the electron energy reaches 4 eV, i.e. the same as the absorption threshold, as shown in Fig. 3. Ultimately the emission always falls when the electron energy is 4 eV or more, but there may be an initial rise. This is characteristic of oxygen arriving at the cathode following decomposition of the barium oxide on the anode. With the alkali halides the behaviour is similar, except that the thresholds for decomposition occur at lower energy than that at which a large absorption effect occurs. Careful observation shows, however, two things:

Firstly, that there is a small absorption effect where decomposition sets in, and secondly, that when the main absorption is reached, the decomposition is intensified. This is seen most clearly with potassium chloride and potassium

iodide. With sodium chloride the poisoning above 4·5 eV is so marked that it is difficult to detect the intensification. In some of the films of barium oxide there is similarly an intensification of the poisoning when 7 eV is reached, where there is another absorption effect, as mentioned above.

We have studied the decomposition of other compounds, giving thresholds as shown in Table 2, though with these we have not made corresponding observations of the absorption effects.

Table 2. *Decomposition and absorption thresholds*

	First threshold for decomposition	*Threshold for main absorption*		
NaCl	4·5 W	7·7 W		7·3 H
KCl	4·0 Ja, W.	6·5 W	7·0 Ja	7·1 H
KI	2·5 W	5·5 W		4·8 H
$BaCl_2$	5 W			
$BaSO_4$	7 W			
FeO	5·5 Ja			
CuO	2·0 Ja			
MoO_3	7·5 Ja			
	8·0 W			
NiO	2·5 Ja, W.			
WO_2	4·5 Ja			
Ta_2O_5	4·6 Ja			
ZrO_2	1·7 Ja			
BaS	1·5 W			
SrS	2·0 W			

Some of these figures are in agreement with those observed by Jacobs.[7] It is worth mentioning that in experiments of quite a different type Tompkins and Jacobs,[8] working on the photolysis of barium and potassium azide, have found reason to believe that excitons are involved, i.e. the excitation processes corresponding with the first main absorption band, and their detailed argument could probably be modified to apply to the alkali halides and to barium oxide, to explain the decomposition process in detail. The decomposition at a lower energy in the alkali halides is more difficult to explain. It may be due to transitions initiated at the surface in the neighbourhood of lattice vacancies or "kinks."[9]

Thus with a wide variety of compounds, decomposition occurs when they are bombarded with electrons of sufficient energy, and we now have some knowledge of the thresholds and of the mechanisms involved. In all cases, the gaseous constituent is liberated, and ultimately poisons the emission of an oxide cathode, although initially very small quantities of either oxygen or halogen may enhance the emission slightly. The decomposition process leaves excess of metallic constituent near the surface of the compound under bombardment, which lowers its secondary emission coefficient.

THE DIFFERENCE BETWEEN D.C. AND PULSED EMISSION FROM A THERMIONIC CATHODE

When the cathode temperature is raised in a diode with a contaminated anode, the emission current density can be increased, causing more rapid gas evolution from the bombarded anode film, but at the same time the cathode becomes more resistant to poisoning. The result is that the threshold cannot be determined, and there may even be none detectable. At high voltage and current density, the poisoning may be very rapid indeed, for example with a time constant of as little as a millisecond. This is a primary cause of the difference between d.c. and pulsed emission. When contaminating films are present on the anode, there is a considerable difference between the two at temperatures where both can be measured, e.g. 800–900° K. Thus rapid measurement of the d.c. emission may give a value less than the pulsed by a factor of 5 or more. Continuous d.c. operation with the cathode below 1000° K may increase this factor by lowering the d.c. emission further. Heating of the anode at 900–1000° C, or a prolonged bombardment with the cathode above 850° C, and anode volts above the decomposition threshold, will remove or decompose many contaminants, in which case the d.c. emission will be stable at the lower temperatures, and the value at 900° K will be not less than a third of the pulsed emission. With a thoroughly outgassed anode, e.g. a tungsten sheet heated at 2700° K for some hours, the d.c. and pulsed emission are equal from 800 to 900° K. It is difficult to measure d.c. emission above 900° K, since the anode becomes over-heated, both liberating gas and causing back-heating of the cathode. Cooling of the anode is of course possible, but a geometry which permits this makes a thorough initial outgassing difficult, and in any case further deposits will form on a cool anode with the cathode above 900° K. Measurements at higher temperatures with a water-cooled anode show that the pulsed-d.c. ratio increases above 900° K.[10] We have been inclined to suppose that this is due to the anode contaminations, increasing during operation, but this has not been completely proved. The alternative suggestion[11] is that there is a difference resulting from movement of ions in the coating, and this remains a possibility at the higher temperatures. At the lower temperatures, however, there is no support for this, and the decomposition poisoning seems able to explain all the observations.

ACKNOWLEDGEMENT

The author desires to tender his acknowledgement to the M-O. Valve Co. Ltd., on whose behalf the work described in this publication was carried out.

REFERENCES

(1) Bell, J. *Proc. Roy. Soc. A.*, **141,** p. 641 (1933).
Bell, J., Davies, J. W., and Gossling, B. S. *J. Inst. Elec. Engrs.* **83,** p. 176 (1938).
Metson, G. H. *Brit. J. Appl. Phys.*, **1,** p. 73 (1950); **2,** p. 48 (1951).
(2) Myers, H. P. *Proc. Roy. Soc., A*, **215,** p. 329 (1952).
(3) Rudberg, E. *Phys. Rev.*, **50,** p. 138 (1936).
(4) Apker, L., Taft, E., and Dickey, J. *Phys. Rev.*, **84,** p. 508 (1951).
(5) Tyler, W. W. *Phys. Rev.*, **76,** p. 1887 (1949).
Taft, E. A., and Dickey, J. E. *Phys. Rev.*, **78,** p. 625 (1950).
Tyler, W. W., and Sproull, R. L. *Phys. Rev.*, **83,** p. 548 (1951).
de Vore, H. B., and Dewdney, J. W. *Phys. Rev.*, **83,** p. 805 (1951).
(6) Hilsch, R. *Z. Phys.*, **77,** p. 427 (1932).
(7) Jacobs, H. *J. Appl. Phys.*, **17,** p. 596 (1946).
Jacobs, H., and Dobischek, D. *Phys. Rev.*, **81,** p. 1019 (1951).
(8) Tompkins, F. C. *Ind. and Eng. Chem.*, **44,** p. 1336 (1952).
Jacobs, P. W. M., and Tompkins, F. C. *Proc. Roy. Soc., A*, **215,** pp. 254, 263 (1952).
(9) Wright, D. A., and Woods, J. *Proc. Phys. Soc. [London], B*, **66,** p. 1073 (1953).
(10) Wright, D. A. *Proc. Phys. Soc. [London], B*, **62,** p. 398 (1949).
(11) Nergaard, L. S. *R.C.A. Review*, **13,** p. 464 (1952).
(12) Lempicki, A. *Proc. Phys. Soc. [London], B*, **66,** p. 278 (1953).
(13) Jonker, J. L. H. *Philips Res. Rep.*, **2,** p. 331 (1947).

New books

The atomization of liquid fuels. By E. GIFFEN and A. MURASZEW. (London: Chapman and Hall Ltd.) Pp. ix + 246. Price 36*s*.

Liquid fuels play an essential part in the modern world and, almost always, efficient atomization is a prerequisite to their effective performance. The authors are therefore to be congratulated on bringing together for the first time into a single volume the results of the widespread research which has been carried out on this important subject. The book deals with basic factors affecting atomization rather than with the engineering design of equipment, and its scope is well illustrated by the chapter headings: (i) Mechanism of disintegration of liquid jets, (ii) Fuel spray characteristics, (iii) Effect of atomizer on spray characteristics, (iv) Simple theory of swirl atomizer, (v) Use of dimensional analysis for correlation of atomization data, (vi) Effect of physical properties of the liquid on spray characteristics, (vii) Effect of the properties of the gaseous medium, (viii) Effect of injection pressure, (ix) Formation and development of intermittent and continuous sprays, (x) Experimental methods for assessment of spray characteristics.

The subject is, on the whole, well presented although there is some repetition between different chapters. This is perhaps inevitable, because the variables under which it is discussed are not independent. Dimensional analysis so far fails to provide a completely satisfactory basis for the correlation of data, but the chapter on this subject illustrates the value of this method of approach. A few detailed criticisms can be made, for the benefit of future editions. The simple arithmetical exercises in Chapter I could well be omitted, and if numerical data must be given in both English and metric units the values should correspond, see p. 9, "2·6 mm (0·12 in.)." Typographical errors in mathematical expressions were noted on pp. 30, 58, 123 and 132. The paragraph on logarithmic probability distributions is unlikely to enlighten anyone. But these are small matters; it is a good book with a good bibliography, well indexed and well produced, which should be on the desk of all interested in the problem of liquid atomization.

W. H. WALTON

Atmospheric electricity. 2nd edition. By B. J. F. SCHONLAND. (London: Methuen and Co. Ltd.) Pp. vii + 95. Price 7*s*. 6*d*.

The subject matter of *Atmospheric electricity* is by convention restricted to the electric fields and currents of the lower atmosphere in fair and disturbed weather; it includes the study of the lightning discharge. It is a narrowly defined subject with comparatively few devotees, and so is admirably suited to treatment in the series "Monographs on Physical Subjects," which requires a "compact statement of the modern position" in about 25 000 words.

The first edition appeared in 1932; the present book retains the same plan and the same concise style, but the chapters on thunderstorm electrification and the lightning flash have been completely rewritten, and the small and carefully chosen bibliography brought up to date. Those who know the first edition will require no recommendation, others can be assured that there is no better introduction to the subject. It will particularly repay the attention of any teacher looking for examples to enliven his presentation of electrostatics.

G. D. ROBINSON

An introduction to relaxation methods. By F. S. SHAW. (New York: Dover Publications Inc.) Pp. 396. Price $5.50.

This is a useful working manual for relaxation techniques applied to a variety of mathematical and physical problems. The development follows the usual lines: linear algebraic simultaneous equations, ordinary differential equations, second- and fourth-order linear partial differential equations, eigenvalue problems, integral equations. A full bibliography of the method is appended. There is obviously some difficulty in communicating what is essentially a mathematical art in a way which makes easy reading for a non-specialist. The author's difficulties are not alleviated by the format, the book size being $8\frac{1}{4} \times 5\frac{1}{2}$ in., which gives a page far too small for a pleasant set-up. There is a curious unevenness in the spacing of the type within words, which is distracting. In general, this book, expensive in dollars, would be a worthwhile addition to a reference library, but it does not appear to provide the type of introduction to the subject sought by the general student.

I. I. BERENBLUT

Foodstuffs, their plasticity, fluidity and consistency. Editor: G. W. SCOTT BLAIR. (Amsterdam: North Holland Publishing Co., 1953.) Pp. xv + 264. Price 50*s*.

Dr. Scott Blair in his introduction to this volume says: "The study of the rheology of foodstuffs is still almost entirely at the empirical stage of development." The contents of this book confirm this view, so that it is of greater interest to the food technologist than to the physicist. Separate chapters are concerned with starch; cereals; milk, cream, ice cream mixes and similar products; butter; cheese and curd; honey; confectionery, gels, sweets, chocolate and marzipan; and a final chapter on the psycho-rheology of foodstuffs. Each chapter is a useful summary of existing knowledge, although, in one or two, a more critical approach would be of assistance.

The chapter by J. Pryce Jones stands out as a model of the physical approach, in which considerable success has been achieved in relating properties and composition.

Physicists concerned with the relation of subjective judgement to instrumental measurement will find many points in the book to interest them. The chapter by Dr. R. Harper provides a full review of the approach developed by Scott Blair, Harper and their co-workers.

ALAN G. WARD

Probability and information theory, with applications to radar. By P. M. WOODWARD, B.A. (London: Pergamon Press Ltd.) Pp. x + 128. Price 21*s*.

During the past few years our whole approach to the study of telecommunication has undergone a radical change. Signals are no longer treated as though they were determinate events, but are viewed in the light of probability theory. Signal analysis, in terms of Fourier-type methods, is now distinguished from statistical communication theory. This theoretical work, which is largely due to Wiener and to C. E. Shannon, sets up a measure of the *information-content* of physical signals, in terms of their statistical rarity, together with a general method of assessing the upper limiting *capacity* of a noisy channel to transmit information.

The inclusion of the expression "information theory" in the title of this book is slightly unfortunate, since this term is

frequently used to refer to certain aspects of scientific method (concerning the "extraction of information from nature," by experiment). There are indeed relations between the two fields, but they are not to be identified.

The reviewer confidently recommends this short book. Wisely, the author starts at the beginning of the subject, and surveys the elements of (statistical) probability theory. The beginner's feet are placed firmly on the ground. An excellent summary of the elements of signal analysis is presented; Fourier analysis, convolution, the sampling theorem, correlation functions, and the ideas of signal and spectral moments, etc. A section on noise is followed by a general treatment of the problem of *detection* in terms of inverse probabilities. Using simple illustrations drawn from radar, the author shows how all empiricism may be removed from this process of detection, as indeed it may be eliminated from other aspects of communication, by this probabilistic approach to the question of signal *information*. COLIN CHERRY

Introduction to electron microscopy. By CECIL E. HALL. (London: McGraw-Hill Publishing Co. Ltd.) Pp. ix + 451. Price 64*s*. 6*d*.

The International Series in Pure and Applied Physics has produced one of its rare contributions to the application of physics, in the form of a textbook on electron microscopy. Professor Hall's account has grown out of the graduate course that he has conducted for some years at M.I.T. for students of biophysics. The audience is thus the user, and especially the biologist, rather than the physicist who is building or operating the instrument. Nevertheless, the fundamentals are in no way sacrificed to descriptive presentation, and a considerable acquaintance with calculus is required for a full understanding of the text.

The introductory chapters explain with thoroughness the nature of the electron and the methods available for calculating or mapping its behaviour in electrostatic and magnetic fields. The properties and construction of ideal and real electron lenses are set out, with a short descriptive chapter on aberrations. Two chapters are devoted to the principles and design of the electron microscope itself, and to a survey of the commercially available models. Before detailing the methods of specimen preparation, naturally the longest chapter of the book, very thorough consideration is given to electron scattering and to the theory of image formation. No such detailed treatment, essential to a complete interpretation of the electron image, has previously been compiled, and it will be found to contain much that is new and valuable to the experienced electron microscopist, as well as to the beginner. The final chapter gives typical examples of applications of the electron microscope, without trying to cover the vast literature that now exists.

The whole of the text is informed with the author's long experience in the field of electron microscopy, as well as by many individual and illuminating ways of presenting physical phenomena to the non-physicist. But, although it is more comprehensive than any other work on the subject, it may be felt that it is too much so. For instance, the nature of the electron is discussed much more fully than it might be thought biologists would require. Either the M.I.T. student is more versatile than his fellows in Europe, or he is not expected to retain more than outline knowledge of the physical principles. On the other hand, it might be held to be important for his understanding of probable future developments of the technique for greater attention to be paid to the theory of aberrations. The neglect of metallurgical applications is not so much the result of the nature of the original lecture course, as of the general backwardness of the United States in using the electron microscope outside biology. It remains to add that the book is admirably produced, with a wealth of diagrams and half-tone illustrations, and with many problems attached to most chapters to test the student's progress. The price is high; how do American graduate students afford it?

V. E. COSSLETT

Communication theory. Edited by W. JACKSON, F.Inst.P. (London: Butterworths Scientific Publications Ltd.) Pp. xii + 532. Price 65*s*.

In September 1952 a symposium on applications of communication theory was convened in London by Professor Willis Jackson. The papers read at this symposium together with the ensuing discussions have now been published.

These thirty-eight papers cover a wide field, and are by no means all theoretical despite the title of the book. The introductory paper by Dr. Gabor is an excellent summary of communication theory intended "as a refresher or compendium for non-specialists" but not "as an elementary introduction." The other papers are divided into groups dealing with transmission systems and coding, signal discrimination, transmission channels, television, speech, hearing and a final section of miscellaneous papers. This last section includes some of the more interesting and original papers and describes work on semantics, generators of information, the structure of language and the optical application of information theory.

This book is certainly neither suitable nor intended as an elementary introduction to communication or information theory. It is a collection of research papers, generally of a very high standard, by authors from both Europe and America. That some of the papers are controversial is evidenced by the often lively discussions but this is to be expected in a volume recording the latest state of knowledge of a subject, and adds much to its interest.

The editing of this collection of papers must have been a most difficult task, but it is disappointing, here and there, to find references in the discussion to unrecorded remarks and to statements in papers which have not found their way into print. The production of the book is most pleasing.

I. L. DAVIES

Dipole moments. By R. J. W. LE FÈVRE. (London: Methuen and Co. Ltd.) Pp. vii + 140. Price 8*s*. 6*d*.

This book covers, essentially, the use of dipole moments as a tool of chemistry. The purely chemical chapters seem useful and cover a fair number of recent publications but, unfortunately, the exposition of the physics of the subject is quite out of date. Dielectric theory is presented as if the theory for gases were all that needs to be known. Dipole moments are measured in gases or dilute solutions, so that it seems reasonable to concentrate on the theory of gases. However, it is hardly possible to justify a treatment in which the reader is not told that for condensed phases dipolar interaction leads to conditions quite different from those applying to isolated dipoles, and that theories for solids and liquids have been developed in recent years. Where some formulae based on modern theories are introduced (in the chapter on "solvent effects in measurement") their general significance is not made clear. It is a curious feature of this revised third edition of a book which first appeared in 1938, that the newly introduced material in the first chapter serves only to make it more antiquated. VERA DANIEL

Notes and comments

Fifth International Spectroscopy Colloquium

The Spectrochemistry and Colorimetry Group of the Austrian Chemical Society under the chairmanship of Professor F. X. Mayer has undertaken the organization of the Fifth International Spectroscopy Colloquium to be held at Gmunden, in the Salzkammergut district of upper Austria. The colloquium will be held from 30 August to 3 September, 1954, and both molecular and emission spectroscopy will be covered, as in previous years. It is suggested that the main subjects for papers should be:

1. Molecular spectroscopy applied to research on artificial fibres; non-dispersive infra-red spectroscopy.

2. Emission spectroscopy applied to the analysis of non-conductors and copper alloys.

It is proposed, as an innovation, to finish the papers one hour earlier and to occupy the hour with short correlated review papers by two speakers, followed by informal discussion of *personal* experiences of the work under review. For such contributions, the following themes are suggested:

Molecular spectroscopy: Critical comparison of photoelectric and photographic method, infra-red and Raman spectroscopy, as analytical techniques.

Emission spectroscopy: Evaluation of spectrograms, experiences with methods not involving standard samples (e.g. Harvey and Addink methods).

Excursions are being planned, and a ladies' programme will be arranged if there is sufficient interest.

The Austrian Group will make the final selection of papers and arrange the programme. Each country is asked to submit not more than five papers, and comments are invited on the above proposals.

The Industrial Spectroscopy Group of The Institute of Physics invites offers of papers, suggestions and comments and would like to have the names of those proposing to attend the colloquium as soon as possible. Papers should be of not more than 15 minutes' duration, and summaries will later be required.

Bibliography of information on servomechanisms and related subjects

The first addendum to the Bibliography of information on servomechanisms and related subjects has now been received. The bibliography was first issued in 1950 and was prepared for the Interdepartmental technical committee on servomechanisms (see *J. Sci. Instrum.*, **25**, pp. 224 and 368, 1948). Limited supplies of the first addendum, the bibliography and some of the report referred to are available from the Ministry of Supply, TPA3/TIB (M), Servo Room 140, Leysdown Road, Mottingham, London, S.E.9.

Analytical Abstracts

We have received a copy of the first issue of a new monthly abstract journal published by W. Heffer and Sons, Ltd., Cambridge, for the Society of Analytical Chemistry and entitled *Analytical Abstracts*. This deals with all branches of analytical chemistry, including techniques and laboratory apparatus. There are 218 abstracts in this first issue. The subscription rate is 50*s.* per annum (inclusive of the index). Orders should be sent to The Society for Analytical Chemistry, 7–8 Idol Lane, London, E.C.3.

Educational booklets on thermionic valves

The first two booklets in a series entitled *Electrons* were issued in 1950 by the Edison Swan Electric Co., Ltd. Another excellent booklet in the series has now been received. The first two dealt with diodes and triodes and this third one covers screened grids and pentodes. The booklets are written in a concise and clear style and their subjects are treated in a thorough manner from first principles. Another three booklets which are planned will deal with beam tetrodes, frequency changers and cathode-ray tubes.

Copies of the booklets are available free of charge to educational authorities and others from the Edison Swan Electric Co., Ltd., 155 Charing Cross Road, London, W.C.2.

Journal of Scientific Instruments

Contents of the March issue

ORIGINAL CONTRIBUTIONS

Papers

A modified Miller time-base circuit. By R. D. Ryan.
A submersible photoelectric absorptiometer. By L. J. Scragg, R. Briggs and G. Knowles.
A balanced electrometer amplifier. By D. R. Hardy.
A simple method for the precision measurement of lattice-constants. By H. J. Goldschmidt.
An electron diffraction apparatus with a new electron optical system designed for the examination of surface structure. By C. S. Lees
Cementing optical glasses to obtain strong joints free from striae. By E. Djurle, E. Ingelstam and L. Johansson.
An apparatus for preparing sodium chloride crystals having large specific surface. By D. M. Young and J. A. Morrison.
An apparatus for producing a number of different gas mixtures for a long period of time. By S. W. Bailey.
A simple thermionic vacuum gauge. By G. K. T. Conn and H. N. Daglish.
A logarithmic chart drive. By P. H. Dawson, F. G. Haynes and M. B. Coyle.
A small masticator for rubber plasticization. By W. F. Watson and D. Wilson.
A computer for solving some problems in connexion with travelling-wave particle accelerators. By M. C. Crowley-Milling.
A photoelectric spectrophotometer for the vacuum ultra-violet region. By V. J. Hammond and W. C. Price.

Laboratory and workshop notes

Galvanometer operated alarm switch. By J. L. Horrell.
Arc welded molybdenum apparatus for high temperature research. By J. W. Tomlinson, D. C. Lowe, G. W. Mellors and J. O'M. Bockris.
A method of verifying Rockwell hardness machine depth indicators. By R. J. Ellis.

NOTES AND NEWS

Correspondence

A method of testing the linearity of response of photomultipliers under pulsed conditions. From W. T. Link and D. Walker.
Some experiments with a cold-cathode vacuum gauge. From K. Phillips.
Notes on the ionization gauge. From L. Riddiford.
The correction formula for the dead time of a Geiger–Müller counter. From J. W. Hughes.

New Instruments, materials and tools
Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2*s.* 6*d.* including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

New advances in the junction transistor*

By E. W. Herold, M.Sc., F.I.R.E., RCA Laboratories Division, Princeton, N.J., U.S.A.

This paper presents a survey of some recent United States' research into the alloy junction transistor. The alloy process is described and analyses are given showing how the current amplification factor, α, varies with geometry, surface recombination and emitter current injection. A simple equivalent circuit is derived, based on the physical constants of the transistor. It is then shown that the chief factors limiting frequency response can be reduced in magnitude in an improved radio-frequency amplifier transistor, which also oscillates at frequencies as high as 75 Mc/s. A battery-operated, portable broadcast receiver uses these transistors to obtain performance equal to or exceeding that of tube receivers.

The point-contact transistor, now 5 years old, seems destined to be supplanted for many applications by a slightly more recent device, the junction transistor.[1] The latter type of amplifier was initially difficult to fabricate but can now be made, by a relatively simple alloy process, in either the *p-n-p* type[2] or the *n-p-n* type.[3] Because such a transistor had properties and geometry which differed from those assumed in available theories, studies were undertaken in this laboratory to clarify the operation and improve the frequency response. The present paper reviews the salient points of these researches as a whole; detailed descriptions of some parts of the work are being published elsewhere.[4,5,6]

THE ALLOY JUNCTION†

The rectifying properties of the *p-n* junction were derived theoretically some years ago.[7] Shockley's application of such junctions to an amplifying device came much later[8]; in the Shockley patent,[1] reference is made to diffusing an impurity into a semi-conductor so as to produce a junction. A diffusion technique for making rectifying junctions was more fully described by Hall and Dunlap,[9] but subsequent publications from the same laboratory indicate that diffusion in the solid state is extremely small.[10,11] Successful methods for making transistor junctions utilize a different technique involving alloying. This has already been recognized[2,3] and is confirmed by experimental evidence to be discussed in the present paper.

Fig. 1 shows the steps which take place in forming a junction by the alloy process, for example, using *n*-type, single-crystal germanium and indium as an impurity. In Fig. 1(*a*) a piece of indium is placed on the germanium. The temperature is raised [Fig. 1(*b*)] and the indium melts. Germanium is soluble in liquid indium, the solubility depending on the temperature. When the temperature is further raised, as in Fig. 1(*c*), enough of the germanium is dissolved in the indium to cause a small depression in the crystal. In principle, the size of this depression does not depend on the time (provided equilibrium is reached in the germanium-indium solution) but only on the area, the volume, and the highest temperature reached by the molten indium. Subsequent lowering of the temperature reduces the solubility of germanium in the indium and the germanium crystallizes out, in indium-contaminated, *p*-type form. *The important factor at this stage is that the nucleation centres come from the original single-crystal base.* As a result, the *p*-type germanium re-crystallizes *in single-crystal form* on top of the base material, and in crystallographic alinement with it. The *p-n* junction forms at approximately the original solid-liquid interface as a flat, abrupt transition of the Schottky type [Fig. 1(*d*)]. The resistivity of the re-crystallized *p*-type germanium is about 0·001 Ωcm, so that it forms an excellent emitter of holes into *n*-type material of 1 to 3 Ωcm.

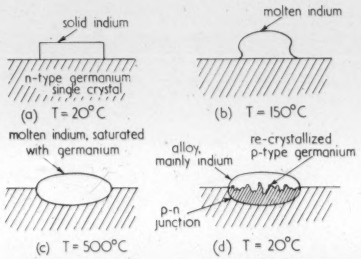


Fig. 1. Steps in the formation of an alloy junction

It is to be emphasized that, in principle, the position and shape of the junction are closely controllable by variation of the contact area, the volume of indium, and the maximum firing temperature. In practice, surface wetting is a possible variant which must also be controlled. Also of note is that any solid-state diffusion will play only a minor part in the junction position, changing it perhaps by a few tens of Ångström units.

The first evidence to indicate the nature of the alloy process was obtained by embedding an actual junction in plastic, cutting it in cross-section and studying it under a microscope. Better results were obtained by first dissolving out the indium with mercury followed by a final nitric acid wash. Some interesting pictures were also made of a rod-like, *n*-type, germanium crystal after it was dipped into molten indium, allowed time for equilibrium, and then cooled to room temperature. Correct choice of crystal direction in the rod permitted observation of re-crystallized germanium on several crystal faces. After removing the indium, the end of the rod looked as in Fig. 2. The typical crystal structure of the base material is clearly evident in the re-crystallized germanium. A view from another angle, i.e. another crystal face, is shown

* Based on a lecture delivered in London on 20 October, 1953.
† Much of the research referred to in this section was done by J. Pankove.

in Fig. 3. X-ray diffraction studies confirmed the basic single-crystal nature of the entire specimen.

After these photographs were taken, the end of the rod was cut in cross-section, etched, and a series of low power

Fig. 2. View of end of single-crystal germanium rod, after recrystallization from an indium melt and removal of the indium

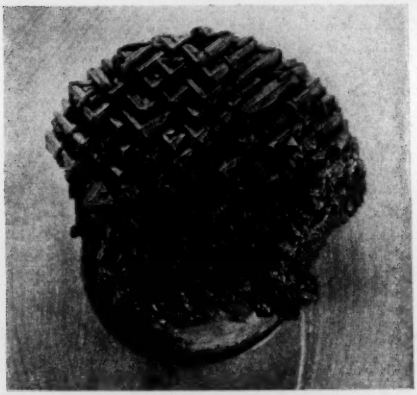


Fig. 3. Same rod as shown in Fig. 2 viewed from another angle

microscope pictures taken, to form the montage of Fig. 4. The various crystal axes are identified by the arrows. By means of thermal probes it was found that a *p-n* junction had formed at a sharp line at about the boundary of the unmelted germanium. This junction line is visible in Fig. 4 and is particularly clear at the 110 faces. The junction is more obscure at the 100 face. Although it may be only a coincidence, it is of interest that this crystal orientation is often avoided by those using alloy transistors because it is said to yield less satisfactory junctions.

THE ALLOY JUNCTION TRANSISTOR

The *p-n-p* alloy junction transistor is made by using two such junctions on opposite sides of a thin germanium wafer. In cross-section, a transistor is shown in Fig. 5. Ordinarily, at least at low frequencies, an input signal is connected to the base lead, the output is taken from the collector, and the emitter, like the cathode of the electron tube, is the common connection. The simplest single criterion of performance applicable to the junction transistor is its value of α, or current-gain factor. This is because any change in emitter current, which fails to reach the collector, contributes to a change in base current and this must be supplied in the form of input power. Ordinarily one thinks of two such current-gain factors, α_{ce}, the ratio of collector-current change to emitter-current change, and α_{cb}, the ratio of collector-current change to base-current change. We want α_{ce} to be unity and α_{cb} to be infinite. It is frequently most convenient to use α_{cb} because it is a more sensitive parameter than the other.

If α_{cb} is to be large, every source of base current must be a minimum. In the main, there are three sources: one is due to imperfect emitter efficiency, a second is bulk recombination of the injected holes in the *n*-type base and the third is surface recombination on the free surfaces of the base material. Let us examine these in turn.

Any part of the emitter current carried by electrons will not be able to arrive at the collector, and so represents base current. Primarily, this electron current is due to the free electrons in the *n*-type base. However, if the base conductivity is very small compared with the emitter conductivity, the emitter efficiency will be high. In the alloy transistor the emitter to base conductivity ratio is of the order of 10^3. Except for the high current density case (e.g. power transistors), which will be examined later, a first approximation allows one to neglect emitter efficiency as a major source of base current.

Appreciable bulk recombination (i.e. short lifetime of minority carriers) in the base material prevents some of the injected holes from arriving at the collector. In most transistor theory, this source of base current is emphasized. However, experience with the low power alloy transistor indicated that bulk lifetimes of 5, 100, or 1000 μsec alike gave equally good results. Since it is a poor sample of germanium indeed whose lifetime does not fall above 5 μsec, one must assume that bulk recombination, again to a first approximation, is not a major source of base current in practical alloy transistors.

We are left with surface recombination to explain the base current in the alloy transistor. Early experiments showed that surface treatment of these units was one of the most crucial and critical parts of the processing technique, and confirm that this source of recombination is a major item in determining α.

Fig. 5 shows why the surface may play such an important role. It is observed that the overall transistor geometry is not parallel plane, as hitherto assumed in most transistor theory. Also one notes that the collector is larger than the emitter. The latter feature was incorporated because it was found to lead to high and more uniform values of α, and it is clear that it results in capturing much of the hole current at

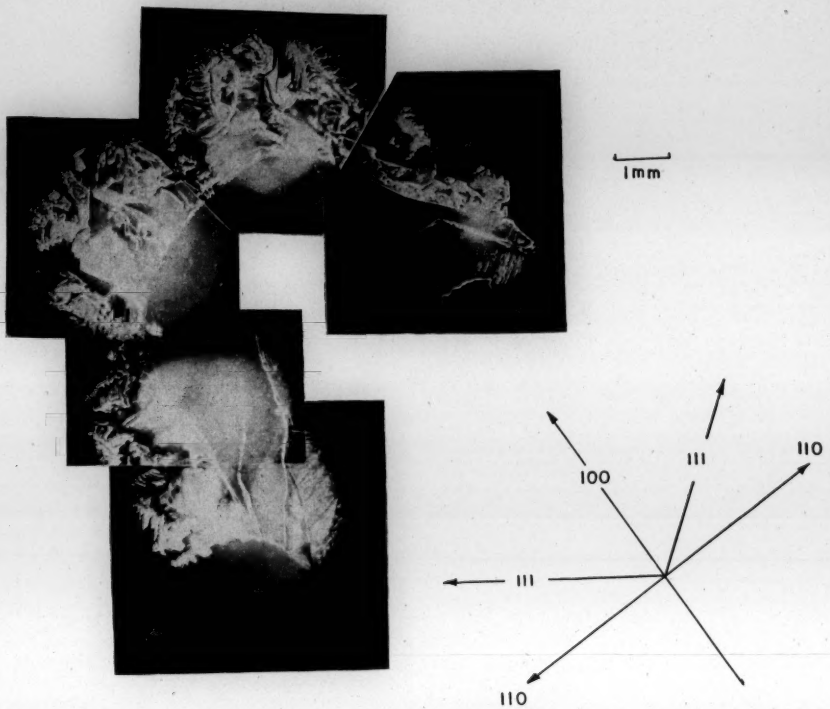


Fig. 4. Montage of cross-section views of end of rod shown in Figs. 2 and 3. The *p-n* junction is seen as a fine white line inside the recrystallized portion

the top of the figure which might otherwise be surface recombination current. It is also seen that there is appreciable surface area surrounding both emitter and collector.

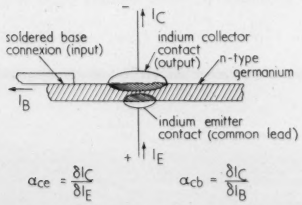


Fig. 5. Cross-section view of a *p-n-p* alloy junction transistor

In the next section of the paper, the special geometry of the alloy transistor will be examined in the light of surface recombination, in order to explain observed values of α. Fortunately, with bulk recombination neglected, all conditions are on the boundaries and powerful analogue methods can be employed to effect a practical solution.

EFFECT OF GEOMETRY ON α*

The minority carriers (holes) flow in the base region primarily as a result of diffusion. The diffusion equation tells us that the current density in the base region is proportional to the gradient of the charge density. Thus

$$I = -qD_p\nabla P \qquad (1)$$

where q is the charge of a "hole," D_p is its diffusion constant, and P is the excess hole density, i.e. the hole density in excess of the thermal equilibrium value. Taking the divergence of each side

$$\text{div } I = -qD_p\nabla^2 P.$$

* The research referred to in this section was carried on by A. R. Moore and J. Pankove.

In the steady state, if there is no bulk recombination, div $I = 0$ throughout the solid, so that

$$\nabla^2 P = 0 \tag{2}$$

a well-known equation for which powerful solution methods are available.

Let us compare this Laplace equation with one similarly derived for a conducting solid of bulk resistivity ρ. If ϕ is the potential.

$$i = -\frac{1}{\rho}\nabla\phi \tag{3}$$

$$\text{div } i = -(1/\rho)\nabla^2\phi = 0 \tag{4}$$

We see, then, an analogy in which $1/\rho$ corresponds with qD_p, and the potential corresponds with excess hole density.

To solve either the transistor or its conducting solid analogue, the boundary conditions must be set up. For the transistor, $P = P_0$, a constant over the emitter electrode surface and, for the usual highly efficient collector surface $P = 0$. In the analogue, corresponding surfaces are connected to a battery which places the emitter analogue at a potential ϕ_0, the collector at potential zero. At the free semi-conductor surface, if there were no surface recombination, there would be no current flow into this surface; the electrical analogue is an insulating surface. When surface recombination takes place, a current density flow proportional to the total charge, qP, takes place

$$I = qPs \tag{5}$$

where s is the proportionality constant, known as the surface recombination velocity. In the electrical analogue, such a surface current can be caused to flow by placing many small electrodes over the entire free surface, and connecting them through resistances back to the zero of potential. If A is the area of each small electrode, and R the value of each resistance, the surface current density is

$$i = \frac{1}{AR}\phi \tag{6}$$

Simple division of the normal component of the general current equations, (1) and (3) by the surface-current equations (5) and (6) shows that

$$\frac{\nabla P_n}{P} = -\frac{s}{D_p} \tag{7}$$

and

$$\frac{\nabla \phi_n}{\phi} = -\frac{\rho}{AR} \tag{8}$$

By making ρ/AR equal to s/D_p the analogy is preserved. We could solve the transistor equation, then, by building a solid conducting model, attaching a battery and measuring the ratio of collector current to emitter or base current.

Fortunately, the thin wafer geometry is such that the powerful tools of two-dimensional analogue solutions, such as the electrolytic tank or conducting sheet, can also be used. Fig. 6 shows the cross-section of a transistor, together with a conducting sheet analogue. If s/D_p is made equal to ρ_s/aR where ρ_s is the surface resistance of the conducting sheet, and a is the length of each free boundary segment, the analogy is approximately applicable. The current flow lines of the analogy are hole-flow paths in the transistor. To take into account the circular nature of the actual transistor surfaces, we may find the current density at each radius in the model and calculate the overall emitter and collector current by weighting each current density value by $2\pi r$. Thus the corrected total emitter current is

$$i_e = \int_0^{r_e} 2\pi r i_e(r)dr$$

and the total collector current

$$i_c = \int_0^{r_c} 2\pi r i_c(r)dr$$

where r_e and r_c are maximum radii of emitter and collector electrodes. The ratio i_c/i_e is α_{ce} so that α_{cb} is readily computed as well.

Although not employed in the present investigation, it is clear that the resistance network analogy board[12] would

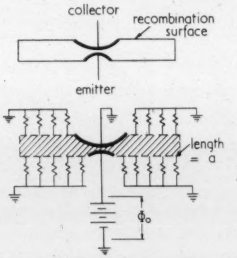


Fig. 6. Two-dimensional analogy to transistor using a conducting sheet. The transistor cross-section is shown above, the analogy arrangement below

permit a direct measurement of the cylindrical geometry; in addition, there is enough flexibility to allow the introduction of bulk recombination effects by use of bleeder resistors plugged in at intermediate points between emitter and collector.

Returning to the resistance-paper solution, Moore and Pankove took data on a variety of geometries and surface conditions.[4] In Fig. 7, a hole-flow map for a typical transistor geometry is shown. One junction is 0·045 in. diameter, the other 0·015 in., and the minimum separation is 0·001 in. Two plots are shown, corresponding to normal and reverse connexions of emitter and collector. A surface recombination velocity of 5000 cm/sec was chosen to correspond with some early transistor data. The surface currents are clearly delineated. The calculated collector-to-base α value was 8 in the normal connexion, but only 0·6 in the reverse connexion.

If there were no surface recombination, the normal and reverse α values would be the same. We arrive, therefore, at a very interesting suggestion. By measuring the ratio of forward-to-reverse α values in an actual transistor, we have a quantitative measure of the surface recombination. The electric analogue solution may be used to supply the calibration curve.

Fig. 8 shows such a calibration curve, calculated for a minimum junction spacing of 0·001 in. (average spacing of 0·0022 in.), and for the typical geometry with one 0·045 in. and one 0·015 in. junction. In practice, transistors vary in both junction spacing as well as surface recombination, and

both factors affect α. However, measurement of the emitter-to-base diffusion capacitance, to be discussed later in this paper, permits an independent evaluation of junction spacing. The forward-to-reverse α ratio, in conjunction with the diffusion capacitance, provides a means of quantitatively

series of early experiments by Pankove in which alloy junction transistors were built with varying geometry. Each point is the average of a number of units. It is seen that highest α value is obtained when the collector area is about twice that of the emitter. The unbroken curve shows the computed

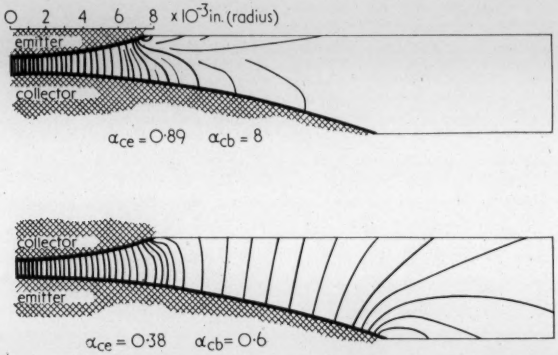


Fig. 7. Hole-flow maps with a surface recombination of 5000 cm/sec. The upper figure shows normal connexion of emitter (0·015 in. diameter) and collector (0·045 in. diameter); the lower figure shows these connexions reversed, i.e. collector diameter 0·015 in., emitter diameter 0·045 in.

determining junction spacing and surface recombination velocity, *even in a finished and already encapsulated transistor*.

Many actual transistors have been measured and it is found that the surface recombination velocity varies from as little as a hundred to many thousands of centimetres per second.

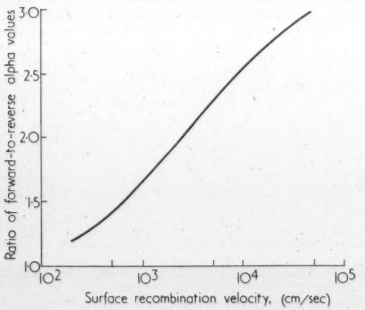


Fig. 8. Calibration of surface recombination for unsymmetrical transistors with 0·015 in. emitter, 0·045 in. collector, 0·001 in. minimum junction spacing and 0·0022 in. average junction spacing

Very instructive experiments have been performed with a given transistor, open to the air, by observing forward-to-back α values while the surface treatment is varied.[4]

Hole-flow maps for different transistor geometries have explained the experimental results which showed the advantage of a large collector area. In Fig. 9, the dotted curve shows a

values using the current-sheet analogue. Reasonable agreement is found.

It should be noted that lack of exact centring of the two junctions, in practice, favours a larger area ratio than the

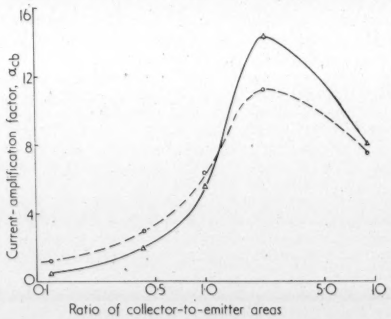


Fig. 9. Comparison of experimental and calculated values of α for *p-n-p* transistor with various area ratios. Largest junction diameter is always 0·045 in. Surface recombination is 5000 cm/sec. Average junction spacing is 0·0022 in.

—— △ —— △ —— calculated values
— — ○ — — ○ — — experimental values

optimum one shown in Fig. 9. A 3 : 1 diameter ratio, therefore, has been widely used. Because modern alloy transistors with a 3 : 1 diameter ratio are found to have surface recombination velocities below 500 cm/sec, the α

values are in the range up to a hundred. The difference over early units is entirely due to improved etching and surface treatment.

EFFECT OF EMITTER CURRENT ON α*

Up to this point, transistor current gain, α, has been discussed under what might be called low power conditions, i.e. the minority carrier charge density in the base region does not greatly exceed the donor concentration. In power transistors, this is not always the case. If the collector-to-base current-amplification factor of a typical *p-n-p* alloy unit is measured, a curve similar to that shown in Fig. 10 is obtained when the

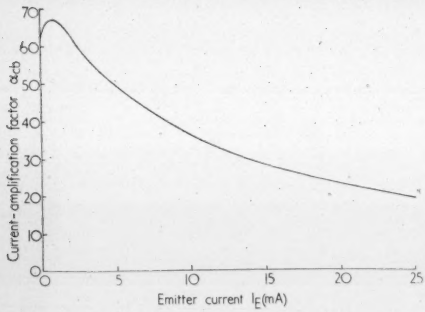


Fig. 10. Curve showing how α varies with emitter current in *p-n-p* alloy junction transistor. Collector diameter, 0·045 in.; emitter diameter, 0·015 in.

emitter current is varied. Low power operation is below a few milliamperes and high α values are found. At high currents, the α value falls off substantially.

It is of interest to derive a relationship for α in terms of emitter efficiency, surface and bulk recombination and then to examine how these terms may vary when the emitter current density is increased. To make the problem amenable to solution, suitable approximations will be made, such as parallel-plane geometry; judicious interpretation makes the results applicable to the alloy transistor as well.

In Fig. 5 a junction transistor is shown. An input variation on the base causes a collector current variation in the load and we are interested in the ratio of the two current variations. The direct currents are I_E, I_B and I_C. The emitter current is comprised of two parts, that due to holes, I_{Ep}, and that due to electrons, I_{Ee},

$$I_E = I_{Ep} + I_{Ee} .$$

With high emitter-to-base conductivity ratio, however

$$I_E \simeq I_{Ep}$$

The base current is the sum of the three factors discussed in an earlier section,

$$I_b = I_{Ee} + I_{VR} + I_{SR}$$

i.e. the emitter electron current, I_{Ee}, the volume recombination current, I_{VR}, and the surface recombination current, I_{SR}.

* The research referred to in this section was undertaken by W. M. Webster.

Since α_{cb} is the derivative of collector current with base current, its reciprocal is

$$\frac{1}{\alpha_{cb}} = \frac{\partial I_B}{\partial I_C} \simeq \frac{\partial I_B}{\partial I_{Ep}} \simeq \frac{\partial I_{Ee}}{\partial I_{Ep}} + \frac{\partial I_{VR}}{\partial I_{Ep}} + \frac{\partial I_{SR}}{\partial I_{Ep}} \tag{9}$$

The first of these terms is related to the ratio of electron to hole current, having to do with emitter efficiency. The second term is the change of volume recombination with emitter current and the third is affected by surface recombination.

Simple first-order theory, for very small density of minority carriers in the base region, gives us the three terms. Shockley and his co-workers[13] show that

$$\frac{\partial I_{Ee}}{\partial I_{Ep}} = \frac{I_{Ee}}{I_{Ep}} = \frac{\sigma_B W}{\sigma_E L_E} \tag{10}$$

where σ_B and σ_E are the base and emitter conductivities, W is the base thickness, and L_E is the diffusion length in the emitter. Also from Shockley and co-workers' paper,[13]

$$\frac{\partial I_{VR}}{\partial I_{Ep}} = \frac{1}{2}\left(\frac{W}{L_B}\right)^2 \tag{11}$$

where L_B is the diffusion length in the base. By a simple integration of the diffusion equation, it can also be shown that[5]

$$\frac{\partial I_{SR}}{\partial I_{Ep}} = \frac{sW}{D_p}\frac{A_s}{A} \tag{12}$$

where A_s/A is the ratio of "effective" recombination surface area to total emitter area and the other symbols are as used previously.

In this paper, it will only be possible to touch upon the highlights of the work and no attempt will be made to complete the derivations. However, we can readily understand the physics of each of the three terms. The first, equation (10), represents simply the ratio of electron-charge carriers to hole-charge carriers available at the emitter junction. The second term, equation (11), is an approximation to the solution of the one-dimensional diffusion equation and gives the loss in hole current because the base thickness is an appreciable fraction of the diffusion length. The final term, equation (12), is proportional to s/D_p, as it should be, and also on the ratio of "effective" surface area, A_s, to emitter area, A, as might be expected.

In the alloy transistor, at low currents, the first term, equation (10), is about 10^{-3} and the second, equation (11), is perhaps a little smaller. The third term, due to surface recombination, is of the order of 10^{-2} and so predominates, as was already seen. However, the present purpose is to re-examine the situation when the direct currents are greatly increased.

Let us see how each term might be expected to vary when the injected hole charge increases to the point where it greatly exceeds the donor density in the base region. Charge neutrality requires an equal number of free electrons to be present, and these are then greatly in excess of the donors. As a result, the electron current through the emitter junction, I_{Ee}, is no longer proportional to the original base conductivity, σ_B, but is much larger. The effect is just the same as if we had apparently increased the base conductivity, i.e. modulated its conductivity. Thus, we see that the first term,

equation (10), increases with emitter current and gives a drop-off of α.

Looking at the second term, equation (11), it is possible that the great excess of free electrons in the base, which accompanies our heavy hole-charge injection, will increase the bulk recombination. This would be the same as a decrease in L_B. If one assumes that "bi-molecular recombination," i.e. recombination proportional not just to the hole density, but to the product of hole and electron density, sets in, the decrease in α due to the second term varies in exactly the same way as that of the first term; both may be thought of as due to an increase in base conductivity. In any event, even if equation (11) is not a strong function of emitter current, the term of equation (10) predominates.

The third term, equation (12), requires a different interpretation. Up to this point in the discussion, no electric field in the base region has been included. All the current which flowed has been assumed to be due to diffusion because of a hole density gradient. When the hole gradient becomes large, however, the electron gradient also becomes large. This gradient would, in the absence of an electric field, cause electron motion in the same direction as the holes. Since electrons cannot escape at the collector, such motion immediately sets up its own restoring force in the form of an electric field, which keeps the electron density high towards the emitter, where the hole density is also a maximum. Such an electric field aids the hole flow toward the collector and, for very high hole injection, gives exactly twice the expected current. It behaves just as if the hole diffusion constant, D_p, had doubled.[5]

Naturally, if the hole flow toward the collector is augmented, that towards the surface is decreased, and the surface recombination is reduced. In fact, as indicated, D_p doubles and surface recombination is halved at high current densities. The term, equation (12), then leads to an *increase* in α with current.

It is now seen how, in the alloy transistor, high emitter currents cause the first two terms of equation (9) to increase in importance, and the last term, originally predominant, to decrease in importance. In place of the low current expression, it can be found that,[5]

$$\frac{1}{\alpha_{cb}} = \left[\frac{\sigma_B W}{\sigma_E L_E} + \frac{1}{2}\left(\frac{W}{L_B}\right)^2\right](1 + kI_E) + \frac{sW}{D_p}\frac{As}{A}g(kI_E) \quad (13)$$

where

$$k = \frac{W\mu_c}{AD_p\sigma_B} \quad (14)$$

and $g(kI_E)$ is a function which varies from unity to one half as kI_E is increased. Fig. 11 shows curves of the function which multiplies the first two terms and that which modifies the last.

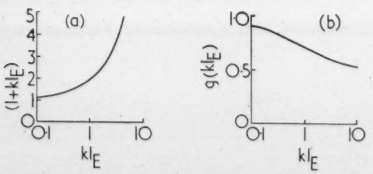


Fig. 11. (*a*) The function which changes the electron-to-hole current ratio and the bulk recombination current when emitter current is increased

(*b*) The function which changes surface recombination current when emitter current is increased

It is well to point out that equation (13) interprets quantitatively each factor in transistor design. The critical nature of W, which enters into every term, even including the factor k, is to be noted. Since the equation is also applicable to *n-p-n* transistors by interchange of subscripts, a comparison of the two is in order. Examining k for germanium, it is clear that μ_e/D_p is about four times as large as μ_p/D_e which applies for the *n-p-n* type. The *n-p-n* transistor is, therefore, considerably less subject to variation of α values with emitter current. This had already been observed experimentally.

By substitution of appropriate numbers into the formula, and the adjustment of the unknown constants so as to obtain

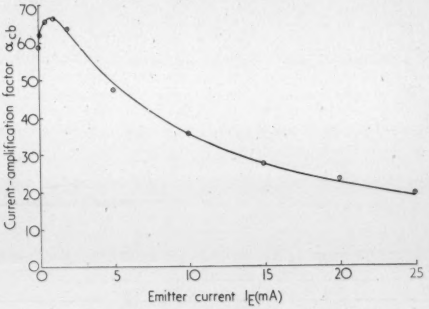


Fig. 12. Comparison of calculated α variation with experiment for typical *p-n-p* alloy junction transistor

——— calculated values
○ experimental values

a reasonable check with surface recombination, Webster made a calculation of α against emitter current for a *p-n-p* alloy transistor.[5] Fig. 12 shows the result, and excellent agreement between the experimental values and the calculated values is found.

Fig. 13 shows experimental and theoretical values for an

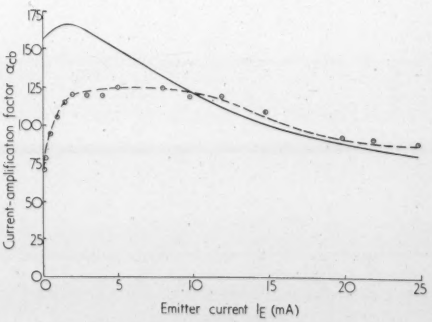


Fig. 13. Comparison of calculated α variation with experiment for typical *n-p-n* alloy junction transistor

——— calculated values
– – ○ – – experimental values

n-p-n alloy transistor. Although agreement is satisfactory for the large current region, the experiments show an even flatter curve than the theory predicts at low currents. The discrepancy is believed to be due to other factors not yet fully investigated. In this connexion, it should be mentioned that less is known of the alloy mechanism in the *n-p-n* type which uses a binary impurity metal.[3] There is also evidence that use of a binary impurity metal for the *p-n-p* type gives a flatter α curve than a single impurity. However, there is little doubt that, qualitatively, the analysis correctly interprets the observed phenomena of decreased α values with increase in current.

HIGH FREQUENCY PERFORMANCE AND THE EQUIVALENT CIRCUIT*

As with other electronic devices, there are high frequency limitations of the alloy junction transistor. A possible limitation may be evident in the hole-flow map, Fig. 7. There are appreciably different path lengths taken by the minority carriers in their travel from curved emitter to curved collector. Such a limitation is not basic, as we shall see, since junctions can be made reasonably parallel by variations in the alloy technique.

Much more basic is the fact that the carriers pass from emitter to collector by a diffusion process. Such current flow is relatively slow and, in addition to an expected phase shift, causes a decrease in α values as the frequency is raised. Much has been written about the so-called "α cut-off" and it is, indeed, a basic limitation in transistors having no electric field to assist flow through the base.

In the simplest, conventional, alloy transistor, however, the decrease in magnitude of α is only partly responsible for the reduced gain at high frequencies. For this reason, it is better to consider an equivalent circuit and then find the various component values by a combination of intuitive and experimental approaches. A first clue was obtained because junction transistor theory failed to predict accurately the static characteristics of alloy junction transistors. In order to investigate the cause of the discrepancy, the transistor admittances were carefully measured.[14] The first one chosen was that of the input, i.e. from the base to the grounded emitter. A bridge was set up and, to observe the degree of balance over a wide frequency range, a square-wave input signal was used as shown in Fig. 14(*a*). A balanced-input oscilloscope was used as null detector, and the transistor was operated with normal bias supplies (collector by-passed to ground). For a typical *p-n-p* alloy transistor, a network which balanced quite well over the wide frequency range of the square wave is shown at Fig. 14(*b*). The large capacitance and its shunting resistance could have been expected from junction theory as we shall see. The 350 Ω series resistance, however, was unexpectedly high and obviously cannot be ignored. Very little intuition was needed to find its source in the transistor construction: the base connexion was made through a substantial length of germanium, $r_{BB'}$, Fig. 14(*c*).

One must recognize that in the equivalent circuit, between the internal intrinsic transistor and the available external base lead, there is a resistor of appreciable magnitude. Once this is recognized, and a correction made, it is found that the *internal* transistor agrees with predicted theoretical characteristics remarkably well, far better than is common with electron tubes. In Fig. 14(*d*) is indicated the external base connexion, *B*, and the internal point, *B'*, which is inaccessible, but which truly represents the base connexion so far as theory is concerned.

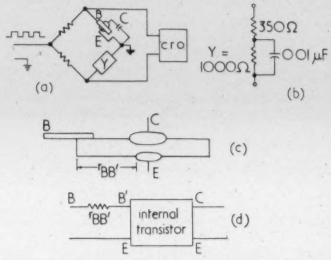


Fig. 14. Investigation of transistor input impedance
(*a*) Bridge for input impedance measurement.
(*b*) Network which balances a typical *p-n-p* transistor.
(*c*) Construction of transistor to show base-lead resistance.
(*d*) Effect of base-lead resistance on equivalent circuit.

We can now draw an equivalent circuit, as in Fig. 15, and begin to evaluate its components from theory. Let us start by considering the emitter junction at such a low frequency that reactive effects are negligible. We know that the current through such a junction, with an applied d.c. voltage V_B, is

$$I_E = I_s\left[\exp\left(\frac{qV_B}{kT}\right) - 1\right] \tag{15}$$

which is the conventional rectifier characteristic and q/kT is a constant of about 39 V^{-1}. For reasonably good transistors, most of this emitter current flows to the collector,

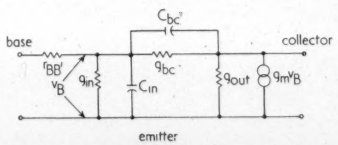


Fig. 15. Equivalent circuit of alloy junction transistor

so that a small input signal will produce a short-circuit output current and

$$g_m = \frac{\partial I_C}{\partial V_B} \simeq \frac{\partial I_E}{\partial V_B} \simeq \frac{q}{kT} I_E \tag{16}$$

and we have then a value for the low frequency part of our output generator. Intuitively, we recognize that the base-to-collector admittance is small, and we can then write down the input conductance, g_{in} as

$$g_{in} = g_m/\alpha_{cb} \tag{17}$$

where α_{cb} is our low frequency current gain. This relation is almost obvious from the definition of α.

At a somewhat higher frequency, reactive effects set in. The output current generator has associated with it the time constant of the diffusion process W^2/D, which leads to both

* The research referred to in this section has been performed by H. Johnson and L. J. Giacoletto.

a time lag and an amplitude loss. In terms of a hypothetical short input pulse, it is delayed, spread out in time, and reduced in amplitude. The solution of the diffusion equation, and use of the approximation that the frequency be not too high, allows us to find an admittance, Y_m, in place of the conductance g_m, which is

$$Y_m = \frac{g_m}{1 + j\omega(W^2/4D)} \tag{18}$$

When a small voltage δV_B is placed between base and ground, a small current flows which charges up the base region. If the voltage is removed, this charge must be swept out again. At low frequencies, the effect is like a capacitance, and quantitatively it is much larger than the normal Schottky-barrier capacitance. The total charge δQ included in the base region is $i\tau$, where τ is the "transit time" or diffusion time. Thus the equivalent capacitance is $\delta Q/\delta V_B$ and is

$$C_{in} = \frac{1}{\delta V_B} i\tau$$

$$= \frac{1}{\delta V_B} g_m(\delta V_B)\tau = \frac{q}{kT} I_E \tau$$

However, the diffusion time τ taken to traverse the base thickness W is of the order W^2/D. The correct solution of the equation, at least for low emitter currents, is:

$$C_{in} = \frac{q}{kT} I_E \frac{W^2}{2D} \tag{19}$$

The formula for C_{in} is well-checked by experiment, so far as variation with I_E is concerned, as may be seen from Fig. 16. At low emitter currents, one finds a linear relation-

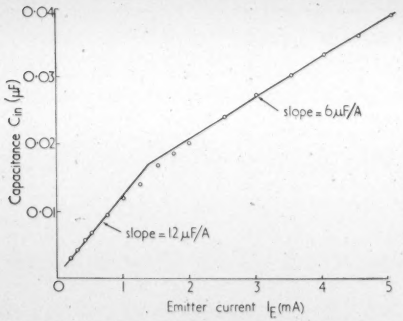


Fig. 16. Change of base-to-emitter capacitance with emitter current showing the two linear regions

ship with one slope, while at higher currents the diffusion constant effectively doubles due to large injected carrier density, as explained above. It is seen that the linear relation again holds, but with half the slope. The experimental evidence, therefore, confirms Webster's predicted effect of an electric field in the base.[5] Because of the direct relation shown in equation (19) between C_{in} and W^2, a measurement of this capacitance is a means of calculating the base thickness W. This is now a standard procedure in this laboratory. We speak knowingly of the base thickness, as if there were a direct measurement and, again, the data can be taken on a completed and encapsulated transistor.

For a base thickness $W = 0{\cdot}002$ in., and a *p-n-p* transistor, the diffusion time is

$$\tau = W^2/2D = 0{\cdot}28\ \mu\text{sec}$$

This seems like a short enough time and, indeed, looking at equation (18) whose time constant has only half this value, this is altered appreciably only when the frequency approaches 1 Mc/s. But C_{in}, for $I_E = 1$ mA, and the same transistor, is calculated (and measured) to be $0{\cdot}011\ \mu$F and, using the measured value of $r_{BB'}$ of Fig. 14(*b*), the time constant of the $C_{in}r_{BB'}$ combination is about 4 μsec. It is concluded that a major frequency effect is due to the input, since C_{in} cannot be compensated by tuning. It is also to be noted that C_{in} is *independent of junction area* and very critically dependent on the base thickness, W.

The equivalent circuit of Fig. 15 uses α_{cb} as a basic low frequency constant for the transistor, as is done with the μ of a triode electron tube. Frequency effects are entirely attributed to constant capacitances and conductances. This has many advantages over the use of a frequency-varying α and it is, perhaps, unfortunate that the latter concept is so prevalent. For the sake of completeness, however, it should be recognized that Fig. 15 can be used to calculate a frequency-varying α_{ce} to be

$$\alpha_{ce} = \frac{(\alpha_{ce})_0}{1 + j\omega(W^2/2D)}$$

in agreement with the accepted formula.

To resume inspection of the equivalent circuit of Fig. 15, a basic cause for g_{out} and g_{bc} is the change in position of the collector barrier when the base-to-collector voltage changes.[15] It is simple to make such a calculation but not too important since surface leakage effects may contribute to these conductances to a significant extent. For the purposes of the present paper, these conductances are small enough to be neglected.

The capacitance C_{out} is that between collector and emitter and is also inherently small enough to be neglected. In fact, diffusion effects can be shown to lead to a small inductance in series with g_{out}, so that, at low frequencies, there is an inherent compensation of C_{out}.

The capacitance C_{bc} cannot be neglected. In the main, it is composed of the barrier capacitance of the collector junction, but there is a small contribution due to the change in barrier position with collector voltage. Of the two significant capacitances this is the only one which depends on junction area.

We shall now discuss the actual circuit values attained in a typical alloy *p-n-p* junction transistor, then show how changes in design can lead to a much improved high frequency performance. Fig. 17 shows the construction, dimensions and the equivalent circuit of a transistor which we shall here designate as the type TA 153 for convenience.[2] A typical unit uses a 0·006 in. base wafer of 3 Ωcm germanium, with a round emitter of 0·015 in. diameter and an opposing collector three times as large. In practice, using pure indium as an alloying element, the junctions have an average spacing, $\overline{W}$, of about 0·0022 in.

At an emitter current of 1 mA, collector voltage of 6 V, a typical α value of 39 is obtained and the 4 kc/s maximum power gain is about 40 db. Such a typical unit has the small-signal circuit constants shown in Fig. 17, all of which can be measured readily. The base lead resistance is of the order of 350 Ω, the diffusion capacitance about 11 000 $\mu\mu$F and the collector junction barrier capacitance (plus its diffusion capacitance) is about 35 $\mu\mu$F. The resistance between

collector and base is an extremely variable quantity but a typical value is 2 MΩ. The current generator has a transconductance of 39 mA/V but, it must be remembered, this is with respect to the internal base-to-emitter voltage, and *not* the applied voltage at the base terminal.

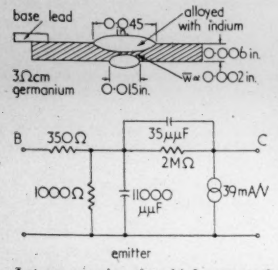


Fig. 17. Typical type TA 153 *p-n-p* alloy junction transistor dimensions and circuit constants at an emitter current of 1 mA

Let us examine the chief causes of poor high frequency performance. First, we would like to reduce the base thickness, W, because this directly reduces the large diffusion capacitance. Then, since the germanium resistance largely determines the base lead resistance, we would like it reduced. Finally, the collector junction barrier capacitance is reduced by cutting down the junction area, so this is desirable. We shall now see how these principles have been applied to a recently developed radio-frequency amplifier transistor.

Before describing this work, perhaps some remarks about other high frequency transistors are in order. The point-contact transistor has been operated as an oscillator up to 400 Mc/s,[16] but point-contact transistors are not well suited for amplifier use because of instability. Other work has been done on "tetrode" junction transistors, and these have been used successfully as amplifiers at frequencies up to the order of 50 Mc/s.[17] Their construction is inherently costly, however, in comparison to the alloy transistors hitherto described.[2,3] The research which we shall now discuss had as its objective the development of an improved radio-frequency transistor by a practical method which would retain the advantages of the alloy technique.

A NEW RADIO-FREQUENCY AMPLIFIER TRANSISTOR*

Fig. 18 shows a cross-sectional view of the new transistor. First, it is noticed that, to reduce the base lead resistance, a relatively thick wafer (0·020 in.) has been used, and the germanium resistivity of the base region is lowered to 0·7 Ωcm. To allow a small base thickness, the collector junction is alloyed inside a small well which is cut into the wafer. This geometry has caused a five-fold decrease in the base-lead resistance. To reduce the collector barrier capacitance, the collector junction diameter has been reduced to one-third. The emitter size was also reduced, to retain the favourable geometry which we have already discussed.

The most important improvement, however, lies in the means used for obtaining reduced base thickness. Because the wafer at the point of alloying is now extremely thin, and the junctions very small, accurate parallelism and control of the alloy penetration is essential. To prevent the relatively deep, hemispherical penetration which occurs when pure indium is used, the indium is first separately alloyed with about 5 atomic per cent of germanium. This partially-saturated alloy, then, can dissolve much less of the single-crystal base and penetrates less deeply than would pure indium. Simultaneously, it leads to a more parallel junction

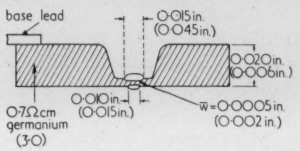
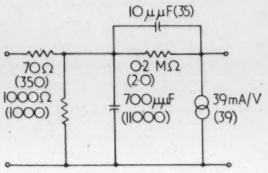


Fig. 18. Cross-section of new *p-n-p* radio-frequency transistor. Dimensions in parentheses are those of the type TA 153, for comparison

surface. The average junction spacing which can be attained in this construction is about 0·0005 in. or less. Since some of the important phenomena are associated with the *square* of this thickness, the improvement in these phenomena is a factor of about 20 or more over the transistor type TA 153 of Fig. 17.

In Fig. 19, we compare the equivalent circuit for the new transistor with values obtained on the earlier type TA 153.

Fig. 19. Comparison of equivalent circuit of new radio-frequency transistor with type TA 153 both at 1 mA emitter current. Figures in parentheses are those of type TA 153



These figures are measured values and, although there are considerable variations among different units, may be considered typical. It is seen that the base-lead resistance is reduced by 5, the barrier capacitance by 3·5 and the diffusion capacitance by about 20 times. The resistance between collector and base, unfortunately, is not as high as that of the type TA 153 transistor. As a consequence, although the value of α remains about the same, the low frequency gain is partially sacrificed. This is not a heavy price to pay for the very great improvement in high frequency performance.

Fig. 20 shows the measured power gain against frequency

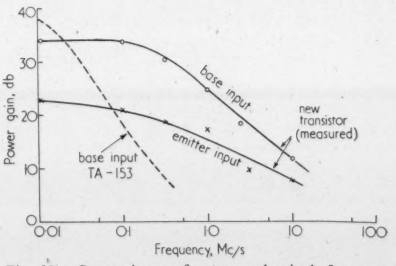


Fig. 20. Comparison of measured single-frequency power gain of a new unit with estimated gain of typical type TA 153 transistor. No neutralization. At I_E = 1 mA, $V_C = -6$ V

* The research referred to in this section has been performed by C. W. Mueller and J. Pankove.

in a simple unneutralized gain test set, in which the output only was conjugate-matched by tuning at each test frequency. The input was a pure resistance. The unit tested was not quite so high in gain at low frequencies as the type TA 153 transistor, but it is seen that appreciable power gain remains (12 db) even at 10 Mc/s. Both emitter-input and base-input curves are included. The former connexion, in analogy to

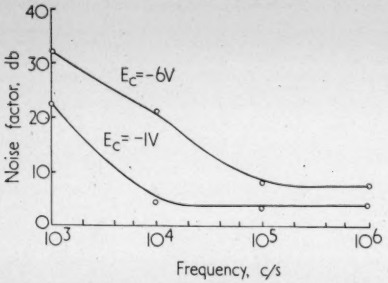


Fig. 21. Comparison of noise factor with frequency for new radio-frequency transistor. Base input. $I_E = 1$ mA

the grounded-grid triode, gives greater stability at the expense of power gain. With base input, maximum gain requires neutralization of feedback and conjugate matching. When this is done, it has been found possible to obtain substantially higher gain than the values measured in Fig. 20 for the frequencies up to 1 Mc/s.

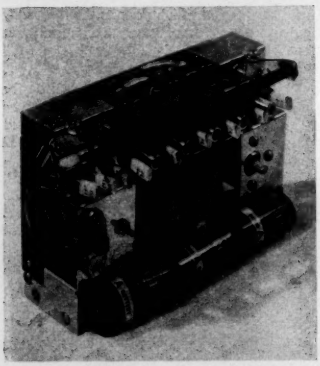


Fig. 22. Photograph of experimental all-transistor broadcast receiver using six of the new radio-frequency transistors, and three other transistors

The noise factor against frequency of one of the new transistors is shown in Fig. 21. It is there seen that the noise factor is quite low for intermediate-frequency or broadcast-band operation.

Although the upper oscillation limit of a transistor, as with a tube, is generally several times higher than its useful frequency range as an amplifier, it is one of the most convenient single numbers to specify performance. Transistors of the new construction described herein have an upper oscillation limit between 40 and 75 Mc/s.

AN EXPERIMENTAL ALL-TRANSISTOR PERSONAL BROADCAST SET*

Although it has been possible to build an all-transistor broadcast set for some time, the performance attained with respect to sensitivity and signal-to-noise ratio was not as good as that of tube receivers. The new transistor just described, however, makes it possible to obtain performance comparable to an all-tube set in the radio-frequency characteristics, and exceed the performance of battery-operated all-tube receivers in nearly every other respect.

Such a receiver has been built and is shown in Fig. 22. It is smaller in size than the battery-operated all-tube receivers being sold in the United States under the designation "personal portable." It has, however, a much larger loudspeaker, about twice the audio output power and yet its low-cost battery of six ordinary flashlight cells will last 500 h.

The receiver has nine transistors and one temperature-compensating junction diode. Six of the transistors are of the new radio-frequency amplifier type we have just described. The remaining three are used for audio frequencies only. The signal is received by a ferrite-cored loop antenna. Miniature intermediate-frequency transformers aid in keeping the chassis size small.

Fig. 23 shows the circuit diagram of the receiver. There is a ferrite-cored, oscillator, capacitor tuned in conjunction with the ferrite loop. The mixer transistor has emitter injection of the oscillator, and base injection of the signal. To provide the correct operating Q, the loop tuner is trans-

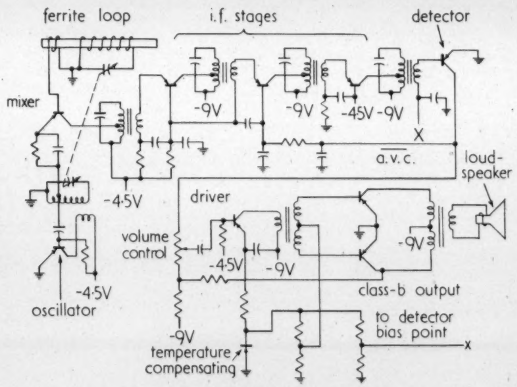


Fig. 23. Circuit diagram of receiver shown in Fig. 22

formed down by a small coupling coil to the base. The three intermediate-frequency stages are operated at 455 kc/s and use the emitter-input circuit. Small, specially-designed, single-tuned, ferrite transformers provide both the selectivity and the proper impedance match for the transistors. There

* The research work referred to in this section has been performed by L. E. Barton.

are both a primary tap and a small secondary winding. A stage gain of about 22 db is achieved.

Automatic volume control (a.v.c.) is obtained from the collector circuit of the detector, which is another of the experimental transistors. Although a.v.c. is shown connected to the base of the second i.f. stage, it will be noted that variations of emitter current of this stage apply variable bias to the first i.f. stage. Thus two stages are gain controlled. The audio system employs a driver stage, transformer coupled to two push-pull class B output transistors. Overall inverse feedback from the output is applied to the low end of the volume control. The 9 V battery supply consists of six medium-size flashlight cells. The supply is centre tapped to permit 4·5 V operation of the r.f. amplifier transistors and to obtain the necessary d.c. conditions for a.v.c. and detection.

The circuits are arranged to reduce temperature variations in the transistors. This is achieved by current-limiting resistances in the emitters, which tend to maintain constant emitter current. The mixer has a self-bias feature which prevents excessive oscillator swing. A special means is provided to give temperature stabilization for the detector and the class B output units. This is done by means of a junction diode, which has a similar temperature characteristic to the collector-base circuit of the transistors. This diode is used in a potential-divider circuit and provides a bias which changes with temperature in the desired manner to maintain proper transistor operation.

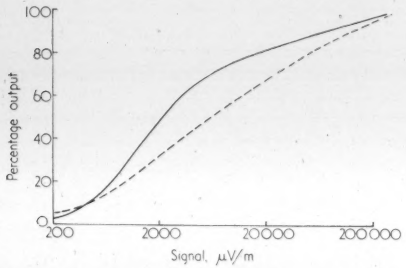


Fig. 24. Comparison of a.v.c. characteristic of transistor receiver with that of a small battery-operated tube receiver

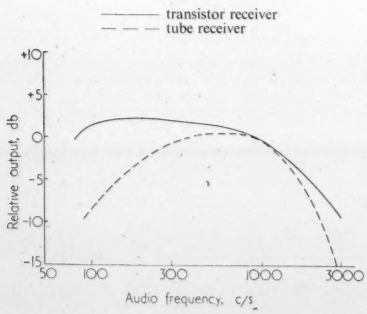


Fig. 25. Comparison of audio response of transistor receiver with small battery-operated tube receiver

The a.v.c. characteristic compared with that of an all-tube receiver is shown in Fig. 24. The audio response curve, similarly compared, is shown in Fig. 25.

Perhaps of greatest significance are the characteristics shown in the table. Of note are the sensitivity and signal-to-noise ratio, which differ insignificantly from the all-tube receiver. The higher audio output, improved quality of the larger loudspeaker, and reduced size and weight of the all-transistor receiver are marked. The really significant advantage, however, lies in the battery consumption, which is over 30 : 1 in favour of the all-transistor receiver.

Comparison of transistor receiver with battery-operated tube receiver

	Tube receiver	*Transistor receiver*
Sensitivity (12·5 mW output), μV/m	125	155
Relative noise, db	0	+4
Maximum audio output, mW	75	150
Loudspeaker size, in.	2×3	4×6
Overall size, in^3	140	73
Overall weight, lb.	3·9	2·7
Total d.c. power from batteries, mW	920	100
Battery life, hours	80–100	500
Battery cost per hour, cents	5·0	0·15

CONCLUSION

The research work of the writer's associates, as summarized herein, has led to a better understanding of the design parameters of the alloy-junction transistor, both for low and high frequency operation. Already this has led to an improved high frequency transistor design which permits the frequency range to be extended by a factor of nearly a hundred over earlier units.

REFERENCES

(1) SHOCKLEY, W. U.S. Patent 2,569,347 applied for 26 June, 1948.
(2) LAW, R. R., MUELLER, C. W., PANKOVE, J. I., and ARMSTRONG, L. D. *Proc. Inst. Radio Engrs*, **40**, p. 1352 (1952).
(3) JENNY, D. A. *Proc. Inst. Radio Engrs*, **41**, p. 1728 (1953).
(4) MOORE, A. R., and PANKOVE, J. *Proc. Inst. Radio Engrs.* In course of publication.
(5) WEBSTER, W. M. *Proc. Inst. Radio Engrs.* In course of publication.
(6) MUELLER, C. W., and PANKOVE, J. *R.C.A. Rev.*, **14**, p. 586 (1953).
(7) DAVYDOV, B. *C.R. Acad. Sci.* [*U.R.S.S.*], **21**, p. 279 (1938).
(8) SHOCKLEY, W. *Bell Syst. Tech. J.*, **28**, p. 435 (1949).
(9) HALL, R. N., and DUNLAP, W. C. *Phys. Rev.*, **80**, p. 467 (1950).
(10) DUNLAP, W. C. *Phys. Rev.*, **86**, p. 615 (1952).
(11) SABY, J. S., and DUNLAP, W. C. *Phys. Rev.*, **90**, p. 630 (1953).
(12) LIEBMANN, G. *Brit. J. Appl. Phys.*, **4**, p. 193 (1953).
(13) SHOCKLEY, W., SPARKS, M., and TEAL, G. K. *Phys. Rev.*, **83**, p. 151 (1951).
(14) GIACOLETTO, L. J. *R.C.A. Rev.*, **14**, p. 269 (1953).
(15) EARLY, J. M. *Proc. Inst. Radio Engrs*, **40**, p. 1401 (1952).
(16) HUNTER, F. L., and SLADE, B. N. *R.C.A. Rev.* In course of publication.
(17) WALLACE, R. L., SCHIMPF, L. G., and DICKTEN, E. *Proc. Inst. Radio Engrs*, **40**, p. 1395 (1952).

The noise of valves*

By C. S. BULL, Ph.D., Electronic Tubes Ltd., High Wycombe, Bucks.

It is shown that a diode valve displays, in addition to the well-known fluctuations in the anode current, independent fluctuations of its characteristics. As a consequence it has become possible to make a less restricted attack on the theory of the fluctuations than has hitherto been done. The new theory is analogous to the application of statistical mechanics to systems which are in thermal equilibrium, but it makes no use of the theorems of thermodynamics. The expressions used take proper account of the fact that the valve is not in thermal equilibrium, and also of the very strong mutual interaction between electrons which is a major feature of the operation of the valve. The resulting equations are valid from low to high anode currents. They are in such a form that the fluctuations can be calculated from the valve characteristic without a knowledge of the cathode temperature, Boltzmann's constant or the separation between the electrodes. It is therefore readily possible to compare the theoretically calculated value of the fluctuations with the observed values, and the extent of the agreement is demonstrated for particular cases.

The effects of screen grids and secondary electron multiplication are calculated by a new approximate method.

During the last century two fields of physics of fundamental importance in the theory of fluctuations developed side by side. One, arising from the invention of heat engines of all sorts, was that of thermodynamics. In this field, matter in macroscopic quantities is treated as having properties expressible by means of variables, the values of which were measured as time averages, and by means of differential coefficients relating these variables one to another. The equations of the subject were often differential equations, and a considerable part of the mathematics of such equations was brought to bear on the subject. Eventually there arose the two laws of thermodynamics which acquired the same universality and completeness as earlier Newton's laws of motion had attained.

At the same time, atomistic views of matter developed, at first in the field of chemistry. Later these ideas extended to all fields of study. For example, in the fields of electricity, light, and mechanics, we have as ultimate units the electron, the quantum and Planck's constant, the unit of action.

Towards the end of the last century, Boltzmann and Maxwell[7,8,9] utilized the atomistic point of view in finding the velocity distributions of gas molecules and in explaining their pressure, thereby connecting the atomistic field with that of thermodynamics through the idea of temperature. In the present century this connexion has been strengthened very greatly, so that starting on a statistical basis, the laws of thermodynamics can be derived. The extreme validity of thermodynamical laws has thus become fully justified.

One of the most important restrictions in statistical thermodynamics is that it deals only with processes in thermal equilibrium. Processes which actually take place cannot, except approximately, fulfil this condition. Another important restriction, which is still not perhaps properly understood and removed, is that the mutual interaction between the particles must contribute negligible energy to the system. Very early, Boltzmann recognized a difficulty at this point, and it is perhaps only recently, in the work of Born and Green,[12] that strong enough methods have been developed to deal with it.

The atomistic view-point almost suggests that fluctuations must occur in all natural phenomena, and that the fluctuations will be of fundamental importance. As early as 1830 the Brownian motion was qualitatively quite correctly explained, but it was not until the beginning of this century that the fluctuations were put quantitatively on a proper basis.

Within the restrictions of statistical thermodynamics, many important fluctuation phenomena have been very successfully dealt with, for example, the blue colour of the sky and the residual fluctuations of galvanometers. For physicists engaged in electronics, an understanding of Nyquist's theory of the fluctuations of resistors in thermal equilibrium is essential, for these fluctuations constitute one of the limitations in the transmission of signals.

The fluctuations of valves impose another limit on the transmission of signals. Now a valve will be found to break through all the restrictions of statistical thermodynamics. It is far from being in thermal equilibrium, for no electron can collide with others sufficiently in the course of its flight through any practicable valve to bring about in any sense the partition of energy described by Maxwell and Boltzmann. On the other hand, the valve often attains its greatest utility only when each electron comes strongly under the influence of many thousands of others in a space charge, its energy cycle and path being very much influenced by them. This interaction can be represented as being due to a conservative field of force, and consequently does not bring about equilibrium. The prestige of statistical thermodynamics, however, is such that some attempts have even been made to use thermodynamical theorems to explain the fluctuations of valves. It is hardly any wonder, therefore, that the subject is at present in a far from satisfactory state, and no two authorities draw the line at the same stage of unreality in developing their views on the subject.

In view of these introductory remarks it is obvious that it would be unwise to accept ideas borrowed from statistical thermodynamics without close scrutiny. We must, therefore, start afresh and at each stage attempt to avoid past errors.

THE PROGRAMME

Our programme will be first to develop a theory of the valve from the macroscopic point of view which will be in keeping with experimental observation. In this, as in thermodynamics, differential coefficients arise. We shall then be compelled to examine the validity of the idea of a differential coefficient as applied to fluctuations. Using the resulting modified values of the differential coefficients we shall develop an equation for the fluctuation current which will be readily comparable with experiment. Finally, curves will be shown demonstrating that the agreement between theory and experiment is satisfactory.

It will not, of course, be possible to develop the theory fully. It has been described completely in detail in four prior

* Based on a lecture delivered on 29 May, 1953, before the Electronics Group at The Institute of Physics Convention in Bournemouth.

publications.[1–4] These publications, however, were written during a period of four years, and some of the ideas we shall study are mentioned in them. This paper, therefore, might be regarded as collecting together the most useful ideas in the earlier papers into a more coherent whole, but we shall also add some new extensions to the theory of the fluctuations.

The purpose of the macroscopic theory is to express the effect of mutual interactions between electrons in the space charge. The examination of the differential coefficients is necessary because, during a fluctuation event, i.e. a time given by $\tau \simeq 1/B$, where B is the bandwidth of the amplifier used, the value of these differential coefficients is in principle not measurable, and it will be shown that their values must be fluctuating. Our work will, therefore, be quite unlike that of prior authors, but will adhere more closely to a sound physical model of the valve.

MACROSCOPIC THEORY

It must be remembered that with d.c. meters we can measure only the macroscopic values of variables for relatively long times, obtaining long-term mean values as a result. For a diode valve the variables may be chosen in a variety of ways, but we shall select I, $\overline{I_s}$, $\overline{V}$ and $\bar{x}$, related by the function

$$I = I(\overline{V}, \overline{I_s}, \bar{x}) \tag{1}$$

$\overline{I_s}$ is chosen instead of cathode temperature since it varies very rapidly indeed with temperature, and may, by this means, be regarded as being very widely adjustable without appreciably changing the emission energies. The range of cathode temperatures over which a valve cathode is useful is, in fact, very restricted, being limited at the lower end by lack of emission and at the upper end by excessive evaporation, leakages, or overheating of other electrodes.[1]

Now we shall divide the valve into N independently fluctuating elements. The area of each element is determined by, and is approximately equal to, the square of the distance from the cathode to the potential minimum. N is usually very large, being about 10^4 for ordinary space charge limited conditions, but is equal to unity for the retarding field condition. Consequently equation (1) must be rewritten

$$\sum_{p=1}^{N} \overline{i_p} = \sum_{p=1}^{N} \overline{i_p}(\overline{V}, \overline{i_{sp}}, \bar{x}) \tag{2}$$

These elements can, to a considerable degree of accuracy, be regarded as fluctuating independently.[2] They are in fact physical realities, which should be taken into account in developing the theory of the fluctuations.

Equation (2) is still, however, not sufficiently detailed to deal with the fluctuations, and we require in addition its time dependence. To express this we divide time into a succession of equal periods $\tau \simeq 1/B$, $\tau_1, \tau_2 \ldots \tau_n \ldots$ etc. such that the instantaneous condition is defined by

$$\sum_{p=1}^{N} i_p(\tau_n) = \sum_{p=1}^{N} i_p[V(\tau_n), i_{sp}(\tau_n), \bar{x}]$$

since x does not vary with time.

For the fluctuations, we are concerned with the behaviour of the increments of this equation at a particular instant, the nth, namely

$$\sum \delta i_p(\tau_n) = \sum \left(\frac{\partial i_p}{\partial V}\right)_{xi_s} (\tau_n) \cdot \delta V(\tau_n) + \sum \left(\frac{\partial i_p}{\partial i_{sp}}\right)_{xV} \partial i_{sp}(\tau_n) \tag{3}$$

The suffixes on the differential coefficients, which will from this point on be omitted, show which variables are kept constant, and we must not confuse the differential $(\partial i_p/\partial V)(\tau_n)$, for example, with the mean value $\overline{(\partial i_p/\partial V)}$. All we know from the earlier work[1] is $\overline{(\partial I/\partial V)}$, or the time averages of the macroscopic values of the differential coefficients. We can also state that $\overline{(\partial I/\partial V)} = N\overline{(\partial i/\partial V)}$ and $\overline{(\partial i/\partial i_s)} = \overline{(\partial I/\partial I_s)}$, if the N elements are all equal.

MICROSCOPIC THEORY

In the last equation we have proceeded from the macroscopic field to the microscopic, both in time and physical size. It is shown in Ref. (3) that, whatever be done, we cannot measure $(\partial I/\partial V)(\tau_n)$. Much less, therefore, can we measure $(\partial i_p/\partial V)(\tau_n)$, and it is clear that the same applies to $(\partial i_p/\partial i_{sp})(\tau_n)$. We must, therefore, either prove that these differentials are constant, or else that they fluctuate. It is quite straightforward to do the latter.

To conserve the energy relations of the cathode we need only assume that the emitted electrons follow, on the average, a Maxwell–Boltzmann distribution corresponding to the cathode temperature. During a particular fluctuation event the distribution could be, and probably is, different from the Maxwell–Boltzmann distribution in a large variety of ways. For example, the electron energy might have been bunched near a value ensuring that more than the usual number of electrons would pass the barrier. The voltage $\delta V(\tau_n)$, therefore, if it were negative during this event, will fail to stop the expected average number of electrons passing. This will constitute a state of affairs in which the slope $(\partial i_p/\partial V)(\tau_n)$ is low, or even perhaps negative. Arguments such as this applied to $(\partial i_p/\partial V)(\tau_n)$ and $(\partial i_p/\partial i_{sp})(\tau_n)$ lead to the conclusion that these differentials, being influenced not only by the fluctuations in the number of electrons, but also by their velocity distribution, fluctuate independently of one another and of the anode current and total emission.

Having determined that the differentials themselves fluctuate, it would be best if we could calculate the extent of the fluctuations. This appears to involve a prohibitive amount of work, so that we are, at least for the present, reduced to putting limits on the fluctuations and making plausible guesses. It is to be hoped that it will eventually become possible to calculate properly the fluctuations of the differential coefficients.

CONSEQUENCES OF THE INSTANTANEOUS INCREMENTAL EQUATION

We can express the instantaneous values of the differentials in equation (3) by the equations

$$(\partial i_p/\partial V)(\tau_n) = \overline{(\partial i/\partial V)}[1 + \epsilon_p(\tau_n)] \tag{4}$$

and

$$(\partial i_p/\partial i_{sp})(\tau_n) = \overline{(\partial i/\partial i_s)}[1 + \nu_p(\tau_n)] \tag{5}$$

It is easy to see, as is shown in Refs. (3) and (4), that the incremental equation (3) using this substitution, when squared and long-time average values are taken, reduces to the following equation for the fluctuation current of a valve.

$$\overline{\Delta I^2} = \left[\left(\overline{\frac{\partial I}{\partial I_s}}\right)^2 (1 + \overline{\nu^2})\overline{\delta I_s^2}\right] \Big/ \left[1 + \left(\frac{R}{R + \frac{1}{\overline{G}}}\right)^2 \cdot \frac{\overline{\epsilon^2}}{N}\right].$$

which by definition of $\overline{(\delta I/\delta I_s)^2}$, reduces to

$$\overline{\Delta I^2} = \left[\left(\frac{\overline{\delta I}}{\delta I_s}\right)^2 \overline{\delta I_s^2}\right] \Bigg/ \left[1 + \left(\frac{R}{R + \frac{1}{\bar{G}}}\right)^2 \cdot \frac{\overline{\epsilon^2}}{N}\right] \tag{6}$$

The denominator of equation (6) is not, as has hitherto been supposed, independent of R, but depends on the mean square deviation of the differential $(\partial I/\partial V)$ or $(\partial i/\partial V)$. If we put $R = 0$, we get

$$\overline{\Delta I^2}_{R=0} = \overline{(\delta I/\delta I_s)^2}\ \overline{\delta I_s^2} \tag{7}$$

an equation obviously of the correct form, once the fluctuation in $(\partial I/\partial I_s)$ is recognized. Hitherto the equation

$$\overline{\Delta I^2} = \overline{(\delta I/\delta I_s)^2}\,\overline{\delta I_s^2} \tag{8}$$

has been used. The fluctuation calculated from this last equation is well known to be far too small if Schottky's equation for $\overline{\delta I_s^2}$ is used, i.e.

$$\overline{\delta I_s^2} = 2e\bar{I}_s B \tag{9}$$

It will also be noticed that $\overline{\Delta I^2}_{R=0}$ must be calculated as the product of two mean square factors, and in consequence we have now greater freedom of theoretical discussion. We had previously only the possibility of attempting to modify equation (9).

ESTIMATES OF $\overline{(\delta i/\delta V)^2}$ AND $\overline{(\delta i/\delta i_s)^2}$

Let us deal first with $\overline{(\delta i/\delta V)^2}$. We can clearly postulate that the mean square fluctuation of the entire valve must be equal to the sum of the mean square fluctuations of the N elements [see appendix of Ref. (3)]. If the valve be regarded as being on open circuit, i.e. $R = \infty$, and the total slope has a deviation at the instant τ_n given by $\epsilon(\tau_n)$, we have

$$\Delta I(\tau_n) = \delta V(\tau_n \,.\, \bar{G}[1 + \epsilon(\tau_n)]$$

Taking mean square values, we find

$$\overline{\Delta I^2} = \overline{\delta V^2} \,.\, \bar{G}^2 \,.\, [(1 + \overline{\epsilon^2})] \tag{10}$$

Now the instantaneous current through an element is

$$\Delta i_p(\tau_n) = \delta V(\tau_n) \,.\, \bar{g}[1 + \epsilon_p(\tau_n)]$$

The mean square current is

$$\overline{\Delta i^2} = \overline{\delta V^2} \,.\, \bar{g}^2(1 + \overline{\epsilon^2}) \tag{11}$$

Therefore the mean square current due to all the elements fluctuating independently is

$$\overline{\Delta I^2} = N \,.\, \overline{\Delta i^2} = \overline{\delta V^2} N \bar{g}^2(1 + \overline{\epsilon^2})$$

$$= \overline{\delta V^2} \,.\, \bar{G}^2\left(\frac{1}{N} + \frac{\overline{\epsilon^2}}{N}\right) \tag{12}$$

Therefore, comparing equations (10) and (12),

$$1 + \overline{\epsilon^2} = (1/N) + (\overline{\epsilon^2}/N)$$

This expression utilizes all the knowledge we have of the fluctuations. We are quite unable to bring to bear on it any thermodynamical theorems relating to the state of equilibrium, such as Nyquist used when considering the fluctuations of resistances.

If we suppose that $\overline{\epsilon^2} \ll 1$, then since N is usually very large, we find

$$\overline{\epsilon^2}/N \simeq 1$$

If we suppose that $\overline{\epsilon^2} > 1$, we find

$$\overline{\epsilon^2} \simeq \overline{\epsilon^2}/N > 1$$

The value of $\overline{\epsilon^2}/N$ must therefore, for ordinary space charge limited conditions, be equal to or greater than unity. Hitherto, when the fluctuations in $(\partial I/\partial V)$ have been neglected, $\overline{\epsilon^2}/N$ has been in effect given the value zero which we see now to be untenable.

For the retarding region of the characteristic, there is no space charge and $N = 1$. Then we find

$$\overline{\epsilon^2} = \overline{\epsilon^2}/N$$

and the fluctuation is not determined by these considerations.

To find the value of $\overline{(\delta I/\delta I_s)^2}$, we can suppose that any fraction from none to all of the fluctuation in $\bar{I}_s$ passes at any instant to the anode. The average fraction is given by $\overline{(\delta I/\delta I_s)}$, the value of which was calculated in Ref. (1). A probability distribution which gives such a fluctuation is the binomial distribution, which is described by the probability generating function:

$$S(z) = (P + Qz)^R$$

where, by definition,

$$P + Q = 1$$
$$\sigma_s^2 = PQR$$
$$\overline{S} = PR$$

From Ref. (1) we find the average value

$$\overline{S} = \left(\frac{\overline{\delta I}}{\delta I_s}\right) = \left(1 + \frac{\overline{xG_x}}{2\bar{I}}\right)\frac{\bar{I}}{\bar{I}_s}$$

where $\overline{G_x} = \overline{(\delta I/\delta x)_{V1_a}}$. It is convenient to put

$$\xi = \overline{xG_x}/2\bar{I}$$

Then

$$\overline{(\delta I/\delta I_s)} = (\bar{I}/\bar{I}_s)(1 + \xi) = PR$$

It is shown in Ref. (2) that $0 < (1 + \xi) < 1$, and also, of course, $0 < \bar{I}/\bar{I}_s < 1$, so that we can solve this equation for P and R by putting either

$$P = \bar{I}/\bar{I}_s; \quad R = (1 + \xi) \tag{13}$$

or

$$P = (1 + \xi); \quad R = \bar{I}/\bar{I}_s \tag{14}$$

In Ref. (4) it is shown that the tenable solution is equation 13, from which

$$\left(\frac{\overline{\delta I}}{\delta I_s}\right)^2 = (\xi + 1)\left(1 + \xi\frac{\bar{I}}{\bar{I}_s}\right)\frac{\bar{I}}{\bar{I}_s}$$

THE METHODS OF COMPARING THEORY AND EXPERIMENT

The space charge smoothing factor is defined by

$$\overline{\Delta I^2} = \Gamma^2 \,.\, 2e\bar{I}B \tag{15}$$

Using the minimum value of $\overline{\epsilon^2}/N \simeq 1$, and the value of $\overline{(\delta I/\delta I_s)^2}$ calculated in the last section we get

$$\overline{\Delta I^2} = \left[(\xi + 1)\left(1 + \xi\frac{I}{I_s}\right)\right] \Big/ \left[1 + \left(\frac{R}{R + \frac{1}{\overline{G}}}\right)^2\right] \cdot 2eIB \qquad (16)$$

so that, by comparison of these two equations, we find

$$\Gamma^2 = \left[(\xi + 1)\left(1 + \xi\frac{I}{I_s}\right)\right] \Big/ \left[1 + \left(\frac{R}{R + \frac{1}{\overline{G}}}\right)^2\right] \qquad (17)$$

Now the measurement of Γ^2 is not practical, for all we can do is to measure the mean square voltage across the valve and its load resistance R. It is shown in Ref. (3) that this is given by

$$\overline{\delta V^2} = \frac{1}{\overline{G}^2} \cdot \left[\left(\overline{\frac{\delta I}{\delta I_s}}\right)^2 \cdot \overline{\delta I_s^2}\right] \Big/ \left[1 + \left(\frac{R + \frac{1}{\overline{G}}}{R}\right)^2\right] \qquad (18)$$

Here we have corrected the equation appearing in Ref. (3) by using the proper long time averages developed in Ref. (4). Equation (18) shows that if $\epsilon^2/N \simeq 1$, then the slope of the curve of $\overline{\delta V^2}$ against $R^2\Big/\left(R + \frac{1}{\overline{G}}\right)^2$ should have twice the value when $\left[R\Big/\left(R + \frac{1}{\overline{G}}\right)\right]^2 = 0$ than when it is equal to unity, i.e. when R changes from zero to infinity.

For experimental verification of the numerator of the equation for Γ^2 we can find the value of $(\xi + 1)$ and I/I_s from the characteristic. To do this we must first remember that our theory applies only if all the elements of a valve are equal. If they are not, some may saturate before others and our equations fail. We should, therefore, not proceed to too great a value of I/I_s, for we cannot be sure that there are no parts of the valve becoming saturated as this ratio increases. The seriousness of this limitation is illustrated by the fact that the bad saturation of an oxide cathode may, in fact, be due to the lack of uniformity of the emission, some areas of the cathode saturating long before others. However, if $\overline{(I/I_s)}$ is small, the effect of the total emission is negligible, and the theory should be applicable even to a non-uniform valve. The value of $(\xi + 1)$ can be obtained from a diode characteristic by means of the formula,[1,4]

$$\left(\overline{\frac{\partial V}{\partial V_m}}\right)_{Ix} = \frac{eI}{kT} \cdot \frac{1}{\overline{G}} \cdot (1 + \xi)$$

or rearranging

$$(\xi + 1) = \left(\overline{\frac{\partial V}{\partial V_m}}\right)_{Ix} \frac{kT\overline{G}}{eI}$$

By drawing two characteristics of the valve at two slightly different cathode temperatures, we can produce two neighbouring characteristics from which all the factors on the right hand side of this last equation can be determined.

EXPERIMENTAL RESULTS

In Figs. 1 and 2 are shown curves of $\overline{\delta V^2}$ against $\left[R\Big/\left(R + \frac{1}{\overline{G}}\right)\right]^2$ for a diode at two values of anode current at which space charge is appreciable. The agreement with our equations is good, the slope varying in the required manner as R goes from $R = 0$ to $R = \infty$.

In Figs. 3 and 4 are shown the characteristics of a diode (6AL5) at two slightly different cathode temperatures, Fig. 3 being drawn on a logarithmic scale and Fig. 4 on linear scale

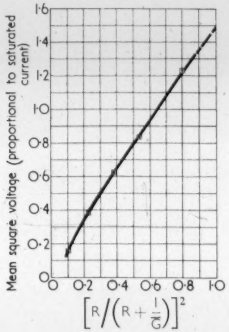


Fig. 1. Curve showing mean square voltage in arbitrary units plotted against $[R/(R + 1/\overline{G})]^2$ for a space-charge-limited current of 1·0 mA

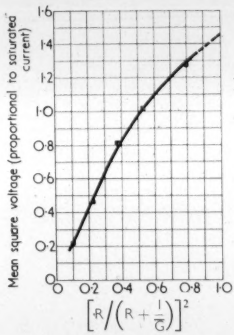


Fig. 2. Similar set of curves to Fig. 1 for a space-charge-limited current of 2·5 mA

of current. The exponential portions of the characteristics aa'' and bb'' are shown extrapolated to a''' and b''' in Fig. 3 and along cc' and dd' in Fig. 4. The equation for $(\xi + 1)$ was evaluated for low anode currents by means of Fig. 3 and for the higher ones by Fig. 4 as indicated in Ref. (4).

In Fig. 5 are plotted the results of the calculation of Γ^2 from Figs. 3 and 4 from very low anode currents up to 20 mA (curve A). For comparison with this curve, curve B is a plot of $kT\overline{G}/eI$, which is often regarded as an approximate value of the smoothing factor. This lies below curve A, and is much lower at the higher currents. Curve C is that displaying the smoothing factor as calculated by D. O. North.[10] It is similar to curve B, but is approximately 1·30 times

higher. North's calculations did not extend into the retarding region.

The experimental points plotted are those obtained in the course of accurate measurements of the smoothing factor of

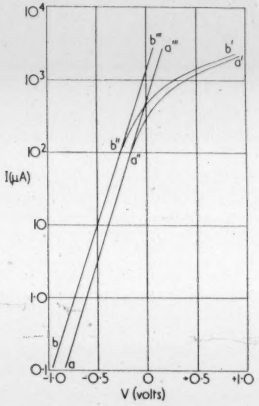


Fig. 3. Characteristics of a diode (6AL5) for low anode currents. Heater voltages 6·3 V and 6·8 V

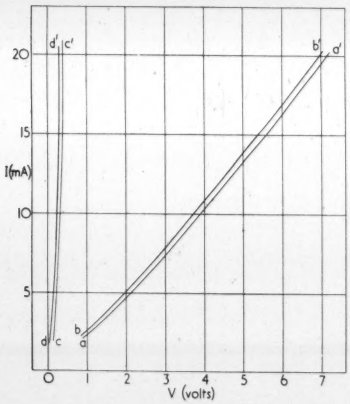


Fig. 4. Characteristics of a diode (6AL5) for high anode currents. Heater voltages 6·3 V and 6·8 V

a similar valve by H. O. Berktay and D. A. Bell.[11] It will be seen that above 10 mA the smoothing factor increases fairly steeply, and this is in conformity with Bell's suggestion that saturation may be setting in. The total emission is, however, much greater than 10 mA, and this effect may be due to the localized saturation mentioned in the last paragraph.

The fluctuations calculated by the new method agree with experimental observations better than do those calculated by earlier methods. The remaining discrepancy between theory and experiment requires closer examination. It is known, for instance, that diodes have much stronger fluctuations than the corresponding triodes. This effect is dependent on the electron optical effects of the grid, and is possibly due to a turbulent flow of the electrons near the potential barrier.

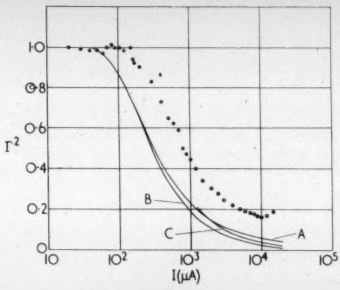


Fig. 5

Curve *A*, $\Gamma^2 = \overline{(\delta V/\delta V_{m1x})^2} kT\bar{G}/e\bar{I}$; curve *B*, $\Gamma^2 = kT\bar{G}/e\bar{I}$; curve *C*, North's results corresponding to curve *B*. Experimental results shown by dots.

The limited amount of experimental results available therefore support the hypotheses made in developing the present theory:

(1) that the properties of the valve fluctuate;

(2) that the mean square deviation of the slope of an element is related to the number of elements by the equation

$$\overline{\epsilon^2}/N \simeq 1$$

(3) that the fluctuation of the total emission is given under all conditions of space charge limitation by Schottky's equation;

(4) that the fraction of fluctuation in the total emission which reaches the anode varies at random from none to all.

FLUCTUATIONS OF TRIODES AND OTHER VALVES

We have been dealing up to this point with diode valves. The fluctuations of triodes, provided the "tail" of the characteristic is not great, can be represented as being due to an equivalent diode, of which the properties are defined in many places,[5,6] and consequently will not be considered in detail here.

Screen grid and more complicated valves, such as pentagrid mixers or thermionic electron multipliers, can be dealt with very straightforwardly by representing the processes taking place in the valve as a sequence of events, each of which changes the electron current by a certain factor. The equivalent diode of such valves must be regarded as being short-circuited, i.e. the anode load $R = 0$, since the effect of screening is that changes in the final anode voltage will have a negligible effect

on the cathode current. For this state of affairs, the instantaneous cathode current is given by

$$\delta I_{R=0}(\tau_n) = (\partial I/\partial I_s)(\tau_n) \,.\, \delta I_s(\tau_n)$$

Now, when the electrons reach a screen grid a fraction F goes on to the anode. This fraction is, on account of the effect of the thermal emission velocities of the electrons, a fluctuating quantity, just as is the ratio $(\partial I/\partial I_s)$. For a screen grid valve, therefore, the anode current is given by

$$\delta I_a(\tau_n) = (\partial I/\partial I_s)(\tau_n) \,.\, F(\tau_n) \,.\, \delta I_s(\tau_n)$$

If this current be allowed to impinge on a secondary cathode with multiplication $M(\tau_n)$, the resulting multiplied anode current will be

$$\delta I_{aM}(\tau_n) = (\partial I/\partial I_s)(\tau_n) \,.\, F(\tau_n) \,.\, M(\tau_n) \,.\, \delta I_s(\tau_n)$$

We now recognize, by analogy with equation (7), that the mean square fluctuation of the product of several independently fluctuating terms is the product of the mean square fluctuation of the terms. We must also remember that just as the fluctuations in $(\partial I/\partial I_s)$ are independent of those of δI_s, so also are those of $F(\tau_n)$ and $M(\tau_n)$.

The fluctuations of the final anode current for the screen grid and multiplier valves are therefore given respectively by

$$\overline{\delta I_a^2} = \overline{(\partial I/\partial I_s)^2} \,.\, \overline{F^2} \,.\, \overline{\delta I_s^2}$$

and

$$\overline{\delta I_{aM}^2} = \overline{(\partial I/\partial I_s)^2} \,.\, \overline{F^2} \,.\, \overline{M^2} \,.\, \overline{\delta I_s^2}$$

The calculation of $\overline{F^2}$ and $\overline{M^2}$ is hardly any more practicable than that of $\overline{(\delta I/\delta I)^2}$. No calculations of the fluctuations in F have been made. The work involved in calculating the fluctuations for a valve having strong electron optical effects, such as an orbital beam valve, is probably prohibitive. The fluctuations in M are also not readily calculable on account of the fact that a clear physical theory of the emission of secondary electrons is not yet available. However, the distribution of $M(\tau_n)$ is nearly a Poisson distribution, so that we can put approximately

$$\overline{M^2} = \overline{M}^2 + \overline{M}$$

FURTHER WORK NEEDED

It is clearly desirable to carry out much more theoretical work on the fluctuations of the characteristic, taking into account the fluctuations in the numbers and the velocities of the electrons.

It must be recognized, too, that a very crude step has been taken in dealing with equations which are valid only for the small fluctuation voltages. Our equations as yet indicate no way in which to put in terms due to signal voltages. Information theory, in which known signals are studied in the presence of noise, will probably be necessary to overcome this difficulty.

The experimental results of Figs. 1 and 2 were obtained between fifteen and sixteen years ago and they may be seriously in error. There is hardly any doubt that modern techniques of measurement and interpretation would enable a much more accurate experimental test of the theory to be made.

ACKNOWLEDGEMENTS

I am very much indebted to my colleagues for very helpful discussion in this subject. In particular, Mr. C. B. Richards has carried out the measurements involved in obtaining Figs. 3, 4 and 5. Mr. J. R. W. Smith has made the suggestion that the way out of the difficulty in applying signal voltages in our equations will be to use information theory.

I am indebted to Mr. D. A. Bell and Mr. H. O. Berktay for permission to use their measurements for comparison with theory. I also wish to record my appreciation of the permission to publish this paper given to me by the Management of Electronic Tubes Limited.

REFERENCES

(1) BULL, C. S. *J. Instn Elect. Engrs*, **97**, Pt. III, No. 47 (1950).

(2) BULL, C. S. *J. Instn Elect. Engrs*, **98**, Pt. III, No. 52 (1951).

(3) BULL, C. S. *Proc. Instn Elect. Engrs* (4), **99**, Monograph No. 41, pp. 289–93 (1952).

(4) BULL, C. S. *Proc. Instn Elect. Engrs* (4), **100**, Monograph No. 61, pp. 47–50 (1953).

(5) KNOLL, M., OLLENDORFF, P., and ROMPE, R. *Gasentladungs-Tabellen*, p. 109 (Berlin: Julius Springer, 1953).

(6) BULL, C. S. *J. Instn Elect. Engrs*, **92**, Pt. II, No. 18 (1945).

(7) MAXWELL, J. C. *Scientific Papers* (2 volumes) (Cambridge: University Press, 1890).

(8) MARGENAU, H., and MURPHY, G. M. *Mathematics of Physics and Chemistry*, pp. 433 *et seq.* (New York: Van Nostrand Co. Inc., 1944).

(9) BOLTZMANN, L. *Vorlesungen über Gastheorie* (Leipzig: J. A. Barth, 1895–98).

(10) NORTH, D. O. *RCA Rev.*, **4**, pp. 269 and 441 (1939–40); **5**, p. 106 (1940–41).

(11) BELL, D. A., and BERKTAY, H. O. Private communication.

(12) BORN, M., and GREEN, H. S. *A General Kinetic Theory of Liquids* (Cambridge: University Press 1949).

The evaluation of interferograms by displacement and stereoscopic methods

By J. W. Gates, M.Sc., A.Inst.P., National Physical Laboratory, Teddington, Middlesex

[*Paper received* 3 *June*, 1953]

A method of obtaining pairs of interferograms is described which makes possible the accurate and rapid evaluation of the optical path variations in the system tested. Such a pair of two-beam interferograms may be measured by displacement methods to an accuracy of about 0·002 wavelength if the path variations are small, or the pair may be viewed stereoscopically to give an impression of the path variation from point to point over the whole area.

The phenomenon of the interference of light is widely applied to the testing of optical systems of all kinds and to various kinds of measurement of length and refractive index. All these applications depend on the way in which the optical path length through the system varies, and on the principle that the pattern of interference bands seen may be taken to represent the change in the optical path from point to point. From photographs of the pattern (interferograms) the variation of path may be calculated and information about the state of the system obtained. Evaluating interferograms is often difficult and tedious, however, and the technique to be described is both more accurate and more rapid than the procedure commonly used. The example chosen to illustrate the method is the comparison of two plane surfaces, but the method is applicable to nearly all examples of interferometry in which high accuracy is demanded.

COMPARISON OF AN IMPERFECT SURFACE WITH A REFERENCE SURFACE

When a nearly-plane surface is set parallel to a plane reference surface as in an interferometer of the type devised by Twyman[1,2] or Fizeau,[3,4] or as described by Tolansky[5] in the case of very small surfaces, the bands which are seen represent height contours referred to a base plane parallel to the perfect mirror or to its image. If the imperfections are several wavelengths in height, the pattern of contour-bands maps out the surface in detail and adjacent bands are sufficiently close for an accurate interpolation of height to be made at any point. If the imperfections are smaller in height, fewer bands will be seen and they will be farther apart. If the variations of height are less than half a wavelength, only one band will be seen and will appear very diffuse. Variations of height will show up as variations in the darkness of the band. Multiple beam interference will show a greater variation of darkness for the same height changes and is therefore more sensitive than two-beam interference, but for the accurate measurement of the contours of such a surface it is convenient to tilt one of the surfaces until a number of parallel bands are seen. Variations of height are then represented as variations in the straightness and in the spacing of the bands in the pattern. This procedure has been described by Guild.[6] The accuracy of the result depends on the accuracy of measurement of the position of a band from point to point, and on the band spacing; the closeness of positions at which measurements can be made, on the other hand, depends on the closeness of the bands of successive orders. These two aims conflict, and increasing the band spacing to improve the accuracy implies that information will be obtainable at fewer places on the surface. By means of appropriate changes of air pressure in the interferometer, Saunders[7] has recorded successive multiple-beam interferograms on the same plate with a shift of a fraction of an order between exposures; in this way the sensitivity associated with a wide band spacing is retained and yet more points on the surface are measurable. Unfortunately, this method is only convenient for small plane surfaces with high reflexion. The present paper indicates how the accuracy of measurement of the position of the bands may be improved so that a closely-spaced band system may be used. This improvement is obtained by simple optical means and does not involve photoelectric measurements as in the method of Werner and Leadon.[8]

LIMITATIONS OF THE STANDARD METHODS

Two main difficulties are encountered in evaluating an interferogram from the straightness of the bands:

(1) it is difficult to estimate the centre of a band accurately, and to set a cross wire accurately on the centre of an image of the band;

(2) variations in the illumination of the surfaces or in the local transmission of the optical parts of the apparatus, or the effect of absorption of light at the reflecting surfaces, may cause a local shading of the bands to one side.

As a result, it is usually found that the accuracy of measurement of a two-beam interferogram is limited to about 0·02 wavelength. The method to be described avoids the first difficulty by using as a reference a band of a similar density distribution, instead of a cross wire; since this reference band is obtained from nearly the same point of another interferogram of the same optical system, the second difficulty also is largely overcome.

METHOD

The matched plane surfaces are arranged as described above with a slight angle between them so that a pattern of about fifteen or twenty nearly straight bands is obtained. Two photographs are taken; in the first the reference surface is tilted to the left: in the other the reference surface is tilted to the right. Care is taken to keep the same spacing and position of the bands in each photograph as far as the imperfections of the surface permit, though the bands increase in order number in opposite directions in the two cases. It has been found useful to fix three small opaque markers on one surface and to adjust the levelling of one surface until one band passes through two of the markers,

and another band of a convenient difference of order (10 or 15) passes through the third marker. When the second interferogram is made, the same apparent configuration is again obtained. Small adjusting devices of high sensitivity have been developed for this purpose.[9] One of these is placed under each foot of the tripod supporting one of the reflecting surfaces, and gives precise height adjustment to 0·01 μ. Specimens of the interferograms obtained by this method are shown in Fig. 1. It will be seen from the photographs that a bulge on one surface will cause the fringes to be displaced to the left in one interferogram and to the right in the other.

Fig. 1. Pair of interferograms produced by the arrangement described. This pair may be viewed in a stereoscope

MEASUREMENT

After development, the relative displacements of the bands at corresponding points of each of the pair of interferograms are measured with the double microscope shown in Fig. 2.

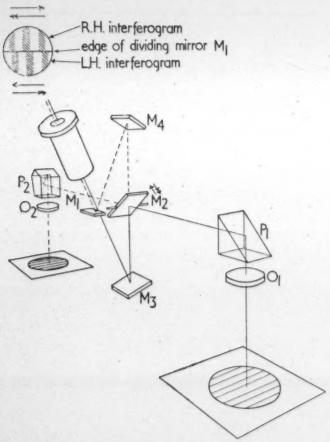


Fig. 2. Double microscope for displacement measurement

The field of the eyepiece of the microscope is divided into two parts horizontally by the mirror M_1; an image of part of one interferogram is presented in one-half by means of lens O_1, right angle prism P_1 and mirrors M_2 and M_3, and an image of the adjacent part of the other interferogram in the other half by lens O_2, pentagonal prism P_2 and mirrors M_4 and M_1. Different prisms are necessary on the two sides for the two images to match.

The images of the bands are arranged to run perpendicular to the dividing edge, and by tilting the inclined mirror M_2 at the centre of the arrangement the two halves of the image may be displaced in opposite directions parallel to the edge. The angular movement of M_2 needed to bring adjacent halves of each band into alinement is measured on a graduated tangent screw, and the movement is equivalent to twice the departure from straightness of the band in each interferogram. This measurement may be converted into height differences by noting the displacement needed between the settings which give alinement of one band in one interferogram with successive bands in the other. The displacement between successive bands corresponds to a height change in the surface of $\frac{1}{4}$ wavelength. If the surface is of reasonably good quality it is found that the calibrating displacement is the same all over the interferogram, so that the movements representing variations in height are proportional to the heights themselves. The microscope is mounted on a linkage

Fig. 3. Contour map from displacement measurements made by the double microscope. Heights in mμ. Contour interval 5 mμ ($\frac{1}{20}$ fringe)

which ensures that the common normal to the axes of the microscope objectives always remains parallel to some fixed line, and the movement of the system is amplified by a pantograph arrangement to facilitate plotting the height measured at a point on the surface at a corresponding point of a map of the surface, and so to produce a system of contour lines (Fig. 3). The standard of "straightness" of the bands in the usual scheme of measurement is replaced in this microscope by the fixed distance between the microscope objectives, and as the microscope is tracked down the pair of images of each band the indicating pointer on the pantograph follows the mean course of the bands. Since the lack of straightness in each band is of the same magnitude but of opposite sign in the two interferograms, the course followed is necessarily straight, and the net of measured loci should consist of parallel equidistant straight lines. By altering the separation of the interferograms the net is moved parallel to itself. This is

equivalent to a small change in height of the plane to which the measurements of height are referred. By rotating each interferogram slightly but in opposite directions a linear change of height along a measured locus may be imposed on the variations of height measured. This is equivalent to tilting the reference plane about an axis parallel to the common normal to the axes of the microscope objectives. By making use of these adjustments, it is possible to make sure that the contour lines produced are in the most useful form for the purpose; for example three selected points may be made to have the same reading. This is equivalent to setting the reference plane to pass through these three points. A tilt of the reference plane about a horizontal axis perpendicular to the common normal to the axes of the microscope objectives is obtained if the spacing of the bands in the two interferograms is not the same. This cannot be corrected in the measuring apparatus, but if the fine adjustment devices are used, sufficiently sensitive settings of the band spacing are possible before photographing the patterns, and undue tilt of this kind is avoided.

ACCURACY

The measurement of height can be made with a sensitivity better than 0·002 wavelength (1 mμ) and with care the results obtained are consistent within this limit. It is a little difficult to decide whether the results have the same accuracy, however, in the absence of some independent method of measuring the contours of the surface to a higher accuracy. It has been noticed, too, that small temperature fluctuations and the like, which for ordinary measurements would be insignificant, may cause local swellings in the surfaces between the photographing of one interferogram and the other. Some experiments have been carried out, however, in which the two interferograms have been obtained in a different way from that described. Between the first photograph and the second the points of support beneath the lower optical flat were moved, but the relative inclination of the surfaces was kept the same. The change in the shape of the pattern of bands then represents the difference in the sagging of the flat, which may be calculated. Results so far indicate that the limit of 0·002 wavelength is not optimistic, and that for such small band movements as are encountered in the sagging of a thick plate, the sensitivity of the arrangement is considerably higher.

STEREOSCOPIC EFFECTS

The pair of interferograms obtained as described shows parallax differences from point to point. If a pair such as that in Fig. 1 is viewed in a suitable stereoscope, the images may be fused, at least over part of the area covered, to give an impression of the relief of the surface. The surface seen resembles the original surface, but the height dimension is greatly exaggerated, possibly as much as a million times, depending on the relative tilt which was used and the magnification. Unfortunately, there is a tendency for the height of ridges which run across the direction of the bands to be estimated differently from ridges which run parallel to the bands, but usually quite a good impression of the shape of the surface may be obtained if the lack of flatness is not too great. This method of presentation should prove of great value to the optical glass-worker.

STEREOSCOPIC MEASUREMENT

The difficulty mentioned in the previous paragraph applies only when heights are estimated from the stereoscopic impression without any mechanical aid in their measurement. If a pair of interferograms is projected on a screen by a pair of projectors fitted with red and green filters and the images are viewed through complementary spectacles, a relief image is built up which may extend behind or in front of the screen and it is possible to measure heights by moving the screen towards or away from the projectors until a suitable mark on the screen appears to lie in the image of the surface. This is the way in which contour detail from air survey photographs is filled in, and under certain circumstances the relief image which is traced is a scale model of the surface photographed. This method has been described by Hart.[10] For various reasons it was found impossible to obtain interferograms which would give an accurate image over the whole area, but contours could be rapidly drawn for small parts of the interferogram.

FURTHER APPLICATIONS

The application of the measuring techniques suggested may be applied to two classes of interferometry where:

either (1) a very simple interferogram may be converted into a pattern of parallel bands by the equivalent of the tilting of one surface as in the method described, and it is possible to repeat the pattern with the reversed tilt without unwanted changes taking place in the interval;

or (2) the pattern of parallel bands changes by amounts which represent the quantity to be measured. The deflexions produced by the sagging of an optical flat in the example mentioned are of this kind. Other examples may be obtained in many diverse fields:

(*a*) the change in the interferogram of a lens system as the field angle is changed. The measurements should yield values for the increment of oblique aberration;

(*b*) measurements of the change of shape of a body under the influence of temperature or loading and the like;

(*c*) the density changes in the airstream in a wind tunnel.

ACKNOWLEDGEMENTS

The work described has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory. The author wishes to make acknowledgement to the Director of Colonial Surveys for the loan of Multiplex and other equipment, and to Mr. V. A. Williams and Mr. J. A. Eden of the Directorate of Colonial Surveys for valuable discussions on methods of measurement. Miss J. Dolphin carried out much of the practical work.

REFERENCES

(1) CANDLER, C. *Modern Interferometers*, p. 135 (London: Hilger and Watts, 1951).
(2) TWYMAN, F. *Phil. Mag.*, **35**, p. 49 (1918).
(3) MASCART, M. E. *Traite d'Optique*, Vol. I, p. 502 (Paris: Gauthier–Villars, 1889).
(4) FIZEAU, H. *Ann. Chim. Phys.* (3) **66**, p. 429 (1862).
(5) TOLANSKY, S. *Multiple Beam Interferometry of Surfaces and Films*, p. 45 (Oxford: Clarendon Press, 1948).
(6) GUILD, J. *Dictionary of Applied Physics*, Vol. IV, p. 145 (London: Macmillan and Co. Ltd., 1923).
(7) SAUNDERS, J. B. *J. Res. Nat. Bur. Stand.*, **47**, p. 148 (1951).
(8) WERNER, F. D., and LEADON, B. M. *Rev. Sci. Instrum.*, **24**, p. 121 (1953).
(9) GATES, J. W. *J. Sci. Instrum.*, **30**, p. 484 (1953).
(10) HART, C. A. *Air Photography Applied to Surveying*, p. 318 (London: Longmans Green and Co. Ltd., 1940).

The mechanism of the resolution of water-in-oil emulsions by electrical treatment

By C. A. R. Pearce, M.Sc., A.M.I.Mech.E., A.M.I.E.E., Fuel Research Station, Greenwich, London, S.E.10

[*Paper received* 30 *September*, 1953]

When water-in-oil emulsions and many other suspensions are placed in a direct or alternating electric field the droplets or particles line up into chains in the direction of the field and, in the case of emulsions, the droplets may then coalesce and the phases separate. The alinement of the droplets is discussed in terms of the forces between conducting spheres in a field and it is suggested that their coalescence is caused by the force due to electrical charges on them or the electrical breakdown of the thin layer of dielectric separating them. According to these views it should be difficult if not impossible to dehydrate fine stable emulsions completely by the electrical method.

1. Introduction and General Survey

As part of the work of this Station on the dehydration of emulsions of sea-water in residual fuel oil a detailed examination is being made of the mechanism of the electrical process of dehydration and this paper records the preliminary phases of this study.

The use of electricity for the dehydration of water-in-oil emulsions is widely practised in the oil industry but no comprehensive explanation of the action seems to have been published. Speed[1] claims that it was F. G. Cottrell, better known for his work on electrostatic dust precipitation, who first suggested the process and he states that Cottrell thought of it as being analogous to that occurring in precipitators. It will be evident later that the two processes differ in several respects.

Fig. 1 is a photomicrograph of a suspension of one part, by volume, of sea-water in nine of fuel oil and is typical of the majority of the emulsions considered in this paper. It will be observed that the diameters of the droplets range from about 30 to less than 5 μ. Dehydration is effected by means of two electrodes immersed in the emulsion and connected to a source of electric power providing a field strength of up to 1000 V/cm. Although d.c. and a.c. will both bring about dehydration, a.c. is generally more effective and convenient. The temperature of the emulsion, certain additives and possibly other non-electrical factors can affect the rate of dehydration but this paper is concerned only with the electrical effect.

Microscope studies of dehydration. The first studies to be described were made with the aid of a microscope and the two cells shown in Fig. 2. These were constructed from perspex sheet and it was thus possible to watch the dehydration proceeding and to record the stages photographically. In the cell shown in Fig. 2(*a*) the sample of emulsion is situated in an electric field which is sensibly parallel and uniform, while the field due to the electrode system of the cell shown in Fig. 2(*b*) is radial and decreases in intensity with increasing distance from the central electrode.

Fig. 3 is a sequence of photomicrographs of an emulsion initially similar to that illustrated by Fig. 1 and obtained with the concentric electrode cell connected to a 50 c/s alternating supply of 78 V r.m.s. Records of an essentially similar character were obtained using d.c. with the same cell and with each type of supply connected to the parallel electrode cell. With each arrangement, observations were made at a number of positions indicated by the points *A*, *B*, and *C* in Fig. 2.

Dehydration is seen to proceed by repeated coalescence of water droplets. Presumably in the bulk process coalescence proceeds to the point at which sedimentation is effective. The specific gravities of the fuel oils studied ranged from 0·914 to 0·985 so that sedimentation was never rapid and often very slow. The process of droplet growth involves two distinct actions which, however, overlap in time and space. The first step towards coalescence is the bringing together of droplets so that they are separated by only a thin film of oil. Lewis[2] has pointed out that something more than this is necessary to effect the coalescence of droplets surrounded by a double layer of the Helmholtz type. The separating film must be disrupted or the droplets forced together, and this is the second step in the process.

Under the conditions obtaining in most of the studies, the coalescence of adjacent droplets commences while the chains are rudimentary, so that by the time the chains can be readily distinguished, most of their constituent droplets have been formed by the coalescence of two or more of the droplets originally present in the emulsion. By the use of a field strength of less than about 50 V/cm in a narrow streak of emulsion, chains are formed by the alinement of droplets almost without any change of size. With this low field strength the rate of coalescence is negligible and the restricted width of the emulsion film promotes a common orientation of the axes of the chains, so facilitating the combination of numbers of short chains into a relatively few long ones; Fig. 4 was obtained in this way.

It would appear, therefore, that dehydration occurs in two stages which can be studied separately, namely chain formation and coalescence, which are discussed in Sections 2 and 3 of this paper.

2. The Formation of Chains of Droplets

General characteristics of chain formation. The characteristics of chain formation are readily appreciated by a study of Fig. 5, which is a photomicrograph of an emulsion treated in the concentric electrode cell with the emulsion film limited to a small area in the centre, i.e. it did not extend to the outer circular electrode. This photomicrograph is not typical of the records obtained. It is, in fact, a sort of idealized picture, but its clarity will facilitate an appreciation of the salient features of the action observed in all cases, namely:

(i) There is no general movement of droplets in the direction of maximum field strength.
(ii) Chains of droplets are formed by movements of single droplets in directions approximately perpendicular to the length of the chains.
(iii) The chains lie in the general direction of maximum field strength.
(iv) Chains do not always start or finish at an electrode, neither are they all complete; sometimes there are considerable gaps.

To the above must be added the information already given that chains are formed with uniform or divergent fields and with alternating or direct electrical supplies.

Review of phenomenon. Before developing a possible explanation of the phenomenon, some of the more obvious forces that act on particles in electric fields and some of the explanations put forward previously by other investigators will be examined in the light of the foregoing observations.

It is clear that as there is no general movement of the droplets along the line of maximum field strength and as the phenomenon is observed with alternating fields, electric charges residing on the droplets cannot account for the observations, nor can those forces which are responsible for cataphoresis.

Since no general movement of droplets into the strongest part of the field is observed, the reaction between the field and the droplets acting as dipoles from dielectric polarization [see expression (3) Appendix I] cannot account for the chains. In Appendix I an estimate is made of the ratio of the forces between adjacent droplets due to the induced dipoles and those arising from the interaction between the field and the dipoles. In showing that, of these forces, the former do not exceed the latter over an important part of the field of view in the concentric electrode cell it is also proved that the interaction between adjacent dipoles is inadequate to account for the chain formation. Some of the explanations published in very general terms are not therefore acceptable, e.g. that of Eddy and Eddy.[3]

Kuczynski,[4] who observed chain formation in emulsions with high frequency alternating fields showed that, from energy considerations, chain formation might be expected. Muth[5] also worked with high frequencies but with emulsions of oil in milk. His explanation is elaborate but interesting. If it were true it would presumably be equally valid for the sea-water droplets in fuel oil. Muth assumed that the droplets were converted into strong dipoles which alternated in sign synchronously with an applied high frequency alternating field. The dipole effect was assumed to be due to the two ends of the droplets protruding alternately through the outer Helmholtz layer as it vibrated synchronously with the alternating field.

Chain formation in other systems. The formation of chains in an electric field is not limited to emulsions but has been observed in many other systems. Thus Winslow[6] observed "fibration" by an electric field in suspensions of semi-conductors in paraffin, and Hollman[7] observed chains of "semi-conductive" colloidal particles formed in the same way. Putilova, Gindin and Moros[8] also observed chains of carbon black in benzene, of aluminium in petrol, and of zinc and barium stearates in petrol and mineral oil.

It seemed that the first step in investigating the mechanism of chain formation should be to establish the general conditions necessary for its occurrence and to this end powders were obtained with grain sizes up to about 40 μ and of the three groups of materials listed in Table 1.

Table 1. *Substances examined for chain formation*

I	II	III
Copper	Magnesite	Perspex
Zinc	Porcelain	Glass
Aluminium		Quartz
Carbon		
Carborundum		

The particles were not in all cases of regular shape. Those of copper and zinc were practically spherical whilst those of magnesite and perspex were very irregular. Each powder was "mixed" with the same type of fuel oil as was used in work on emulsions in the proportion of about one part by volume of powder to nine parts of oil. In some cases sedimentation was very rapid. A thin layer of the suspension was spread between the electrodes in the cells shown in Fig. 2 and, except that the cover plate was omitted, the effect of connecting the electrodes to a supply of electric power was observed in the same way as in the dehydration experiments.

The photomicrograph shown in Fig. 6 is typical of the effect observed with the first group of materials which fall into the general classification of "good conductors." The materials in the third group, which are all good insulators, did not form chains. The only action observed occurred with the concentric electrode arrangement and strong fields when some particles within a range of about 50 μ of the central electrode tended to move towards it, i.e. into the strongest part of the field. This accords with expression (3), Appendix I. Fig. 7 illustrates the difference between the two actions; it was obtained from a mixture of copper dust and ground glass in fuel oil. The particles of copper are opaque and those of glass almost transparent. The effect of the field on the glass is hardly discernible in this instance. The second group of materials in Table 1 exhibits what may, perhaps, be the property of chain formation, in some small degree, in very strong fields. It was observed that occasionally a few of the particles of these materials which collected around the central electrode in the concentric electrode cell tended to form short radial projections. While these materials come within the general class of insulators they are somewhat inferior to those of the third group, as shown by the values given in Table 2, the inferiority being, in the case of porcelain, associated with relatively low surface resistivity.

Table 2. *Electrical properties of the materials examined*

Material	*$\log_{10}$ (volume resistivity, Ω-cm)*	*$\log_{10}$ (surface resistivity, Ω-cm)*	*Permittivity*
Quartz	14–15	14	4·5
Glass	14	—	5–7
Perspex	17	12	3·6
Magnesite	8	—	—
Porcelain	14	8	4–7

The volume resistivity of the fuel oil used in these experiments was about 10^{10} Ω-cm and its permittivity at 50 c/s was 2·6.

The permittivity of water is about 80 while a good conductor behaves for many purposes as though its permittivity were very high. From Appendix I, equation (1), it will be observed that the effect of the permittivity of a particle on the force acting on it due to an electric field is represented by $(\epsilon_1 - \epsilon_2)/(\epsilon_1 + 2\epsilon_2)$ in which ϵ_1 is the permittivity of the particle and ϵ_2 is that of the medium in which it is immersed. The value of this expression for a conducting particle in fuel oil is almost 1·0, for water in fuel oil it is 0·91, and for glass in fuel oil 0·29. Since the electric field strengths used in these studies were varied over a range greater than 10 : 1 the differences between these values are not significant in relation to the difference in the rates of chain formation for glass and water, and it is clear that the critical property is not simply the permittivity. It would appear that the effect is more directly determined by the resistivity.

Particles of conductors and insulators in a series of liquids were examined in a similar manner to that already described for fuel oil and invariably the conductors formed chains

and the insulators did not. No significant difference was observed in the rates of chain formation after allowance had been made for the viscosities of the liquids. Table 3 gives the resistivities of the liquids that were used.

Table 3. *Resistivity of liquids used in making suspensions*

Liquid	Log_{10} (*volume resistivity, Ω-cm*) *approx.*
Insulating oil	14
Commercial paraffin	12
Deasphalted fuel oil	12
Untreated fuel oil	10
Solution, 4%, of calcium naphthenate in white spirit	8

The last of these liquids was chosen only because, except for mercury, it had the lowest resistivity of the non-aqueous liquids immediately available. The deasphalted oil was very viscous at room temperature on account of separated wax, and chain formation only occurred when it was warm. Distilled water with a resistivity of 10^4 Ω-cm was also tried but electrolytic decomposition prevented reasonable observation and no chain formations were detected.

Suggested mechanism. That chain formation would appear to occur with all conducting particles immersed in a liquid of relatively high resistivity and subjected to an electric field is an indication of the mechanism by which the action proceeds. Groups of such particles distort the equipotential lines and in so doing provide the basis for the necessary forces.

Fig. 8 represents three spherical conductors in a medium of finite resistivity in an electric field. The equipotential lines were plotted by the usual method of an electrolytic analogue. As is to be expected, flow of current is concentrated between the adjacent surfaces of the spheres. In Appendix II an estimate is made of the force on a sphere such as *B* in Fig. 8, from a consideration of the capacitance between the adjacent spheres, i.e. by regarding the adjacent surfaces of the spheres as the plates of condensers. The direction of the resultant force on *B* is clearly such as to move it into the gap between *A* and *C* and it is shown to be about 10^{-5} or 10^{-6} dyn on a sphere of diameter 20 μ, with the arrangement shown in Fig. 9(*b*) and under conditions typical of the experiments, and about 10^{-6} or 10^{-7} dyn when the centre of *B* is 100 μ from those of *A* and *C*, as in Fig. 9(*a*).

It remains to be seen whether these values are of the right order to account for the droplet movements. Observations during the progress of dehydration suggest that a movement of 100 μ in 1 to 5 sec is typical. For a 20 μ diameter droplet the viscous force is the significant factor, the momentum changes being much smaller. The oil used had a viscosity of about 40 centistokes at 20° C and applying the usual expression for the viscous force, namely: $6\pi\eta rv$, where r is the radius of the sphere, η the viscosity of the oil, and v the velocity, a value of approximately 10^{-5} dyn is obtained.

It must be expected that the field strength will frequently exceed that used for the calculation in Appendix II. Two factors will operate to this end. One is the converging field. The other originates in the distortion of the electric field produced by the presence of the chains of conducting particles in the insulating medium and is probably the more important. Once a chain of large droplets is partly formed, the conducting path which they provide acts as a local short circuit and the potential differences across the gaps between groups of droplets become disproportionate to their length. In the light of the above there is excellent agreement between the requirements of the observed movements of the droplets and the magnitude of the forces to be expected from the suggested mechanism.

A little consideration will also show that the type of force postulated fits well with the general characteristics of chain formation, but it should be mentioned that the force of attraction between *A* and *C* which is also present in the arrangement shown in Fig. 8 is likely to be neutralized by the attractions exercised by other droplets and in particular any that have already moved into the line of the embryo chain.

It will be observed from Appendix II that the forces involved in this explanation are dependent on E^2, where E is the electric field strength, and are therefore present equally with alternating and direct fields; further, that these forces are dependent on the separation of charges on the droplets. The axes of the resulting dipoles are not, however, invariably parallel to the field as are those considered in Appendix I but are variously orientated, e.g. that on *B* in Fig. 8 is perpendicular to the field.

3. COALESCENCE

General. As already mentioned, Lewis[2] has drawn attention to the difference between the conditions existing when droplets adhere to one another and those obtaining when coalescence occurs. This section of the paper is devoted to a consideration of the forces which cause droplets in contact, or almost in contact, to coalesce.

A phenomenon which may be only a large scale analogy, but is, nevertheless, instructive, is illustrated in Fig. 10. This shows photographs of two stages of an experiment in which a row of drops of brine in lubricating oil was carefully placed in position between electrodes and then left for some minutes to ensure that the arrangement was reasonably stable. Immediately upon applying a voltage to the electrodes, coalescence occurred.

There would seem to be two factors which might produce coalescence in this way. Since the drops are conductors the sum of the potential differences across the oil films between the drops is equal to the voltage between the electrodes. The oil films are relatively thin and it is to be expected, easily disrupted. It may be, therefore, that coalescence is due to puncturing of the oil layers between adjacent drops. The alternative, or perhaps complementary, factor is the force of attraction between adjacent drops due to their potential difference. This is the force that has already been suggested as responsible for chain formation. The two effects will now be considered separately.

Electrical breakdown of the film between droplets as a factor in coalescence. If the first factor mentioned in the preceding paragraph is of importance in producing coalescence of droplets in electrical dehydration then presumably the voltage across the oil films between droplets is the critical quantity. For a chain of droplets of uniform size (radius r cm) and each almost touching its neighbours, this voltage is $2Er$. When the droplets are of unequal size and r_1 and r_2 are the radii of the two droplets considered, the voltage between them is given by $E(r_1 + r_2)$, to a first approximation.

It will be observed that both the voltage between such droplets and the force between them decrease with decreasing radius. If, therefore, the thickness of the separating film is independent of the droplet size it would follow that fine emulsions should be more difficult to dehydrate than coarse emulsions and this is in accord with experience.

To proceed from this general examination to a quantitative

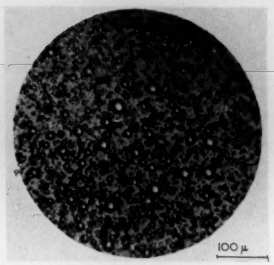


Fig. 1. Photomicrograph of an emulsion of 1 volume of sea-water in 9 volumes of fuel oil

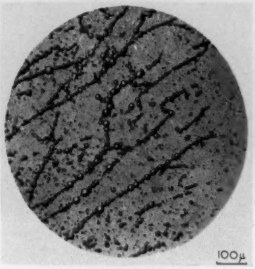


Fig. 4. Photomicrograph of chains of droplets formed without appreciable coalescence

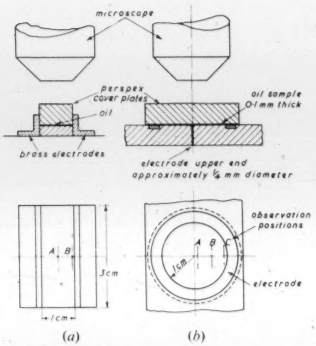


Fig. 2. Microscope cells used for observation of dehydration of emulsions
(*a*) parallel electrodes,
(*b*) concentric electrodes.

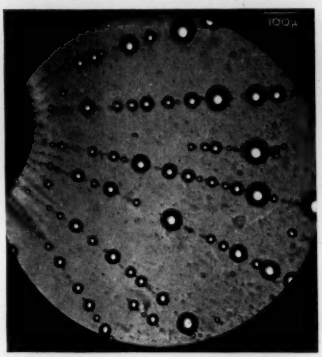
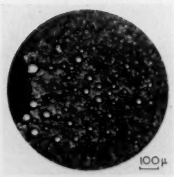
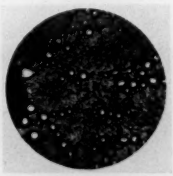


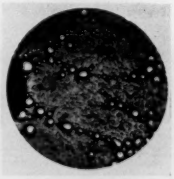
Fig. 5. Photomicrograph of chains formed in an emulsion during dehydration and illustrative of the salient features of chain formation



(*a*) 1 sec after switching on the current



(*b*) after treatment for 1 min



(*c*) after treatment for 3 min

Fig. 3. Sequence of photomicrographs taken while dehydration of an emulsion was in progress. The silhouette on the left is the central electrode of the concentric cell

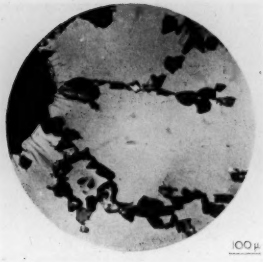


Fig. 6. Photomicrograph of chain formation in a suspension of carborundum particles in fuel oil

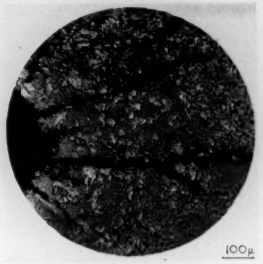


Fig. 7. Photomicrograph of conducting and insulating particles in fuel oil showing the contrasting effects of the electric field. The dark particles are copper and the transparent particles are glass. The large silhouette on the left is the central electrode of the concentric electrode cell

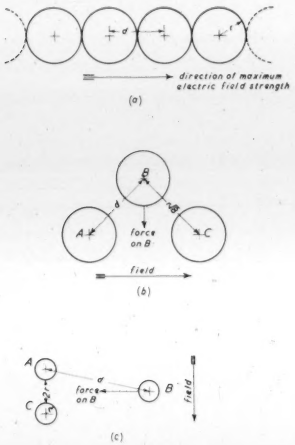


Fig. 9. Arrangements of water droplets in a field
(*a*) continuous chain,
(*b*) droplets out of line,
(*c*) droplets out of line.

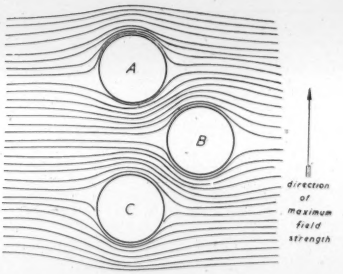
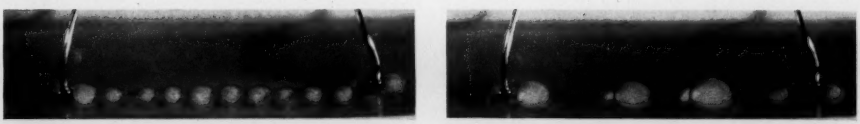


Fig. 8. Equipotential lines, plotted by the electrolytic analogue method, for three spherical conductors in a medium of finite resistivity



(*a*) before switching on

(*b*) immediately after switching on

Fig. 10. A large-scale demonstration of the coalescence of a chain of water drops. Magnification $\times 2\frac{1}{2}$

assessment requires a knowledge of two factors, namely the thickness and dielectric strength of the film separating droplets, immediately before they coalesce. Referring to droplets of oil in an electrolyte, Lewis[2] suggests that they "can approach one another without encountering repulsions until the portions of their surfaces closest together are separated by 10^{-6} cm," i.e. by 2δ where δ is the combined thickness of the Helmholtz layer and the diffuse layer of Guoy. A test of the oil in bulk has given a dielectric strength of 15 000 V/cm which is undoubtedly much greater than the strength of a layer 10^{-6} cm thick. However, if for want of better values these are used to calculate the minimum size of equal droplets that will coalesce in an electric field of strength 500 V/cm due to electric breakdown of the oil film between them, a value of 0·3 μ diameter is obtained.

The force between droplets as a factor in coalescence. The magnitude of the force between spherical droplets of radius r is considered in Appendix II and equation (6) gives the force, in dynes, for droplets in contact as

$$F = 0{\cdot}195 \times 10^{-4} \times E^2 r^2, \qquad (1)$$

which for the fuel oil used in this work and a field of 500 V/cm becomes

$$F = 1{\cdot}27 \times 10^{-7} \times r^2, \qquad (2)$$

where r is in microns.

It is impossible for the spherical surfaces of the droplets to come into contact before coalescence occurs, but the effect of a separation of the order of 10^{-6} cm on the magnitude of the force between droplets of 20 μ diameter is negligible.

Lewis[2] estimated the energy necessary to disrupt the Helmholtz double layers surrounding the adjacent oil droplets in an emulsion of oil in dilute acid as $2{\cdot}9 \times 10^{-11}$ ergs. The assumptions made in the derivation are such as to make the value indicative of no more than the probable order of magnitude. With emulsions of sea-water in fuel oil the interface between the phases is generally assumed to include some asphaltenes and their effect on the interfacial structure has not been investigated. In effect, Lewis's value is equal to the product of a supposed area of contact between droplets and the interfacial energy and there is no reason to assume radically different factors in the present case. For want of a more authoritative value the energy necessary to disrupt the layer separating two adjacent droplets is therefore assumed to be of the order of 10^{-10} or 10^{-11} ergs. If it is further assumed that this energy is equal to the work done by the force between adjacent droplets, an estimate can be made of the minimum diameter of droplets which will coalesce.

The work done in producing coalescence by the electric force of attraction between the droplets is thus $1{\cdot}27 \times 10^{-7} r^2 2\delta$ ergs. As already mentioned, the value given by Lewis[2] for 2δ is approximately 10^{-6} cm, whence the diameter of the smallest droplets that will coalesce may be calculated as being of the order of 18 μ.

Comparison of estimates with observations. No tests were made specifically to determine the smallest droplets that could be made to coalesce electrically but when emulsions with nearly all droplets less than about 3 to 5 μ diameter were examined in the microscope cells no coalescence could be detected. The droplets were observed to form chains under the action of the electric field but the chains were not confined to single rows of droplets as they were with the coarser emulsions; instead they consisted of less regular arrangements three or four droplets wide. In many studies droplets as small as 5 μ were seen to coalesce, but in these instances there were present many larger droplets and it is in accord with the theories presented that it should require a higher field strength to coalesce two droplets each of 5 μ diameter than will suffice with one droplet of 5 μ diameter and one of, say, 15.

Neither of the estimates in the preceding paragraphs is in agreement with the general observation recorded above, but in view of the assumptions the divergence is not surprising. The simplifications imposed on the practical case to render the elementary calculations possible are also responsible for discrepancies; for example, the value of 500 V/cm for the field strength takes no account of the convergence of the field in the case of the concentric electrode system nor the distortion due to chains of droplets. Both these factors increase the field strength in the practical case.

It is of interest to note that if instead of the value 10^{-6} cm suggested by Lewis[2] for 2δ, 10^{-5} cm is substituted, both methods of arriving at the diameter of the smallest equal droplets which will coalesce in a field of 500 V/cm give values according well with the observations, i.e. of about 5 μ.

4. SUMMARY

The work described has suggested a mechanism by which the electrical treatment of water-in-oil emulsions brings about the separation of the water. It has been confirmed that the action proceeds in two steps, as previously observed, namely:

(i) Alinement of the water droplets into chains.
(ii) Coalescence of adjacent droplets until sedimentation becomes effective.

The chains of droplets are formed by forces which originate in the potential differences between the droplets. The presence of conducting particles in a medium of relatively high resistivity produces disturbances in the equipotential lines and the charges induced on the surfaces of the droplets as a result give rise to forces of the correct type and magnitude to account for the observations.

Two effects are suggested as possible mechanisms for bringing about the coalescence of droplets already alined as a chain. The first is the same type of attraction between the droplets as that which was responsible for the chain formation and the second is the electrical breakdown of the oil film between the droplets.

Certain limitations of the process as a means of emulsion breaking follow from this theory, notably:

(i) Droplets smaller than a minimum diameter which is dependent on the conditions, will not coalesce to form larger droplets.
(ii) Partly for this reason and partly because in an emulsion some droplets will not participate in chain formation, it is to be expected that it will be difficult if not impossible completely to dehydrate fine, stable emulsions by the electrical method.

A further practical effect which follows from the theory is that with many oils dehydration will be possible with insulated electrodes. This has the advantage of preventing "flash-over." With a liquid of permittivity 2·6 the displacement current at 50 c/s exceeds the conductance current unless the resistivity is less than about $1{\cdot}5 \times 10^{10}$ Ω-cm.

5. ACKNOWLEDGEMENTS

The work described was part of a programme of the Fuel Research Board of the Department of Scientific and Indus-

trial Research and the paper is published by permission of the Director of Fuel Research.

The illustrations are Crown copyright and are reproduced by permission of the Controller, H.M. Stationery Office.

In conclusion the author gratefully acknowledges the assistance afforded by Messrs. A. H. Perkin and H. A. Howe who made many of the observations, Mr. G. R. Hawley who produced the photographs, and Dr. L. R. B. Shackleton who provided the references to the literature.

REFERENCES

(1) SPEED, B. *Industr. Engng Chem.*, **11**, p. 153 (1919).
(2) LEWIS, W. C. M. *Trans Faraday Soc.*, **30**, p. 958 (1934).
(3) EDDY, W. G., and EDDY, H. C. *Industr. Engng Chem.*, **13**, p. 1016 (1921).
(4) KUCZYNSKI, T. *Przem. Chem.*, **13**, p. 161 (1929)
(5) MUTH, E. *Kolloidzschr.*, **41**, p. 97 (1927).
(6) WINSLOW, W. M. *J. Appl. Phys.*, **20**, p. 1137 (1949).
(7) HOLLMANN, H. E. *J. Appl. Phys.*, **21**, p. 402 (1950).
(8) PUTILOVA, I. N., GINDIN, L. G., and MOROS, L. M. *Dokl. Akad. Nauk SSSR*, **71**, p. 81 (1950).
(9) ABRAHAM, M., and BECKER, R. *Classical Theory of Electricity and Magnetism* (London: Blackie and Son, Ltd., 1950).
(10) ATTWOOD, S. S. *Electric and Magnetic Fields*, 3rd ed. (London, Chapman and Hall, Ltd., 1949).

APPENDIX I

Examination of the forces on droplets in a convergent electric field due to induced dipoles

Abraham and Becker[9] show that the moment induced on a sphere of radius r and permittivity ϵ_1 immersed in a medium of permittivity ϵ_2

$$= E_0 r^3(\epsilon_1 - \epsilon_2)/(\epsilon_1 + 2\epsilon_2) \quad (1)$$

where E_0 is the field strength in which the sphere is situated.

When the sphere is a conductor, ϵ_1 is infinite and the expression becomes $E_0 r^3$.

Assuming that the poles are spaced $2r$ apart their strength becomes $E_0 r^2/2$, and if the three spheres indicated in Fig. 9(*b*) are regarded as dipoles of this sort the resultant force on B can be shown to be approximately

$$\frac{0{\cdot}2}{r^2\epsilon_2} \times \left(\frac{E_0 r^2}{2}\right)^2$$

The permittivity of the oil used in most of the experiments described in this paper was 2·6. Inserting this, the force on sphere B becomes approximately

$$0{\cdot}02\, E_0^2 r^2 \quad (2)$$

A spherical conductor in a non-uniform field is also subject to a force in the direction of increasing field strength (see also Abraham and Becker) which is equal to

$$r^3 E(dE/dx) \quad (3)$$

In a strictly radial field $E(dE/dx) = E_0^2/x$,

and the force $= r^3 E_0^2/x$, (4)

where E_0 is the field at the centre of the sphere and x the distance between the centre of the sphere and the centre of the field.

The force represented by expression (4) exceeds that given by expression (2) for all values of $x < 50\,r$.

Expression (2) is based on the droplets being very close and, even so, for droplets of diameter 20 μ expression (4) exceeds expression (2) for distances up to about 500 μ.

It is therefore reasonable to conclude that since on no occasion did the droplets move into the strongest part of the radial field, i.e. towards the central electrode, the force given by expression (4) was ineffective, and equally that due to the interaction between induced dipoles and given by expression (2) was not responsible for their movements in the formation of chains.

APPENDIX II

Determination of the force between two droplets of water immersed in oil and in an electric field by regarding the adjacent surfaces of the droplets as the plates of a condenser

Attwood[10] shows that the capacity, C, between two equal spheres of radius r immersed in a medium of permittivity ϵ with centres distance d apart is given by the expression:

$$C = 2\pi\epsilon r\left(1 + \frac{r}{d} + \frac{r^2}{d^2} + \frac{r^3}{d^3} + \frac{2r^4}{d^4} + \frac{3r^5}{d^5} + \frac{4r^6}{d^6} + \ldots\right). \quad (1)$$

By summation of the series this becomes

$$C = 2\pi\epsilon r\left[1 + \frac{r}{d} + \frac{r^2}{d^2} + \frac{r^3}{d(d-r)^2}\right]. \quad (2)$$

From the usual energy considerations the mechanical force, F, between such spheres connected to a potential difference V is given by $V^2/2.dC/dd$. Hence F the magnitude of the force, is given by:

$$F = V^2\pi\epsilon\left[\frac{r^2}{d^2}\left(1 + \frac{2r}{d}\right) + \frac{r^4}{d^2(d-r)^2} + \frac{2r^4}{d(d-r)^3}\right] \quad (3)$$

and employing $f(r/d)$ for the expression in brackets

$$F = V^2\pi\epsilon f(r/d). \quad (4)$$

There are two arrangements of spheres which are of interest in the present study and to which the above derivation can be applied. The first case is that of a row of spherical conductors in a liquid of finite resistivity and known permittivity in an electric field, the direction of maximum strength of which coincides with the common axis of the spheres [see Fig. 9(*a*)]. The adjacent spheres are assumed, at this stage, to be separated by a film of a thickness negligible in relation to their radius. The circumstances approximate to a chain of droplets formed between the parallel electrodes of the microscope cell. If E is the field strength, then the potential difference between adjacent spheres is $2Er$, whence if the effect of the remaining spheres on the validity of equation (4) is ignored the force between an adjacent pair is given by

$$F = 4\pi E^2\epsilon r^2 f(r/d). \quad (5)$$

Substituting for $f(r/d)$ with F in dynes, E in V/cm, and r in cm

$$F = 0{\cdot}195\epsilon E^2 r^2 \times 10^{-4}. \quad (6)$$

For the oil used in most of the work ϵ was 2·6 and for a field of 500 V/cm and with r in microns

$$F = 1{\cdot}27\, r^2 \times 10^{-7}. \quad (7)$$

The second condition of interest in this work is the arrange-

ment of spheres represented by Fig. 9(*b*) and (*c*). To determine the order of magnitude of the force on *B* the components due to *A* and *C* respectively are considered separately and assumed to be independent.

Applying expression (5) above, the force on a sphere of diameter 20 μ, situated between two similar spheres as shown, may be calculated for the same general conditions as before to be $0{\cdot}36 \times 10^{-5}$ dyn. If the centre of sphere *B* is distant 100 μ from the common axis of *A* and *C* the force is $0{\cdot}18 \times 10^{-6}$ dyn.

A note on an improved planimetric method of harmonic analysis

By J. CREASE, B.A., and M. J. TUCKER, B.Sc., A.Inst.P., National Institute of Oceanography, Wormley, Nr. Godalming, Surrey

[*Paper first received* 23 *July, and in final form* 8 *December*, 1953]

This note describes an improvement to a method of harmonic analysis described by F. B. Daniels. Both the original and improved methods use only a planimeter and a set of ruled grids. An approximate value of a sine or cosine coefficient is obtained from a single tracing of the plotted function and the accuracy may be improved by means of correction terms from the higher-order harmonics. In the improved method described here the number and importance of the correction terms is much less than in Daniels's original method.

The method described below is an extension of that given by F. B. Daniels,* and in many cases will result in a saving of time. It is most suitable for the harmonic analysis of strictly periodic waveforms, such as engine indicator diagrams, the sustained note from a musical instrument, or the output of an electronic oscillator or of an amplifier fed with a periodic input. It can be conveniently used up to about the fifteenth harmonic. It cannot, in general, be used for non-periodic waveforms, such as tidal curves (which have components whose periods are not rationally related), unless it is permissible to consider a certain length of the curve as one period of a periodic waveform: for certain purposes, for example, a 25-hour tidal curve may be considered as one period of a periodic waveform without introducing serious errors. It is very suitable for casual use because it requires no equipment other than an ordinary planimeter, an instrument readily available in most physics laboratories.

In normal Fourier analysis, the amplitudes a_n and b_n of the sine and cosine components of the n^{th} harmonic of a periodic waveform $f(l)$ are obtained by performing the following mathematical processes:

$$a_n = 2/L \int_0^L f(l) \sin (2\pi nl/L) dl \qquad (1)$$

$$b_n = 2/L \int_0^L f(l) \cos (2\pi nl/L) dl \qquad (2)$$

where L is the fundamental wavelength of the waveform. That is, the waveform being analysed is multiplied by a sinusoidal wave of unit amplitude whose period and phase are the same as those of the component being measured, and the area under the product waveform is then measured over the length of the fundamental period.

In Daniels's method, multiplication is by a square wave instead of a sinusoidal wave, and the product waveform is then integrated as before. This process has the advantage that it can be performed in one simple planimeter operation, and the errors introduced may be corrected from the values obtained for the higher-order harmonics. The process is illustrated in Fig. 1(*a*), where the second harmonic is being measured. The multiplying waveform, or "modulating function," $F(l)$ is shown in Fig. 1(*b*). The process consists of measuring the area under the curve in the regions where the value of the modulating function is $+1$, and subtracting from it the area under the curve in the regions where the modulating function is -1. To achieve this, the planimeter follows the path shown in the outward direction, and returns to the origin along the curve when $F(l) = -1$ and along the axis when $F(l) = +1$. The measurements are performed with the aid of a set of grids ruled on transparent paper or cloth, each grid being moved a quarter of its wavelength sideways for the measurement of the cosine component.

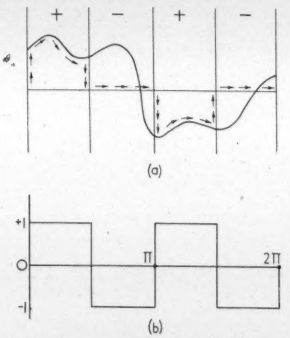


Fig. 1. (*a*) Daniels's diagram showing outward path followed by the planimeter. The homeward path follows the curve in the negative regions and the axis in the positive regions

(*b*) The equivalent modulating function

$$A_n = 4(a_n + a_{3n}/3 + a_{5n}/5 + a_{7n}/7 \ldots)$$
$$n = 2$$

* DANIELS, F. B. *Rev. Sci. Instrum.*, **23**, p. 369 (1952).

For simplicity put $x = 2\pi l/L$ (i.e. the fundamental period becomes 2π). If $F_n(x)$ is the modulating function with n periods in the fundamental period, and A_n is the area measured in the estimation of the sine component, then

$$A_n = \int_0^{2\pi} f(x)F_n(x)dx. \tag{3}$$

Now $F_n(x)$ may be expressed as a well-known Fourier series:

$$F_n(x) = (4/\pi)[\sin nx + (\sin 3nx)/3 + (\sin 5nx)/5 + \ldots]; \tag{4}$$

substituting this in equation (3),

$$A_n = (4/\pi)\int_0^{2\pi} f(x)[\sin nx + (\sin 3nx)/3 + (\sin 5nx)/5 + \ldots]$$

$$= 4(a_n + a_{3n}/3 + a_{5n}/5 + \ldots). \tag{5}$$

Thus it can be seen that A_n contains a contribution from the $3n$, $5n$, etc., harmonics.

By considering the expressions for A_{3n}, A_{5n}, etc.

$$a_n = 1/4(A_n - A_{3n}/3 - A_{5n}/5 \ldots). \tag{6}$$

A similar expression is obtained for b_n.

In the majority of waveforms met with in practice, the amplitude of the harmonic components decreases rapidly with increasing order and it will often be necessary to use only two or three of the correction terms in equation (6).

Daniels did not regard his method from the point of view just set out, and therefore did not realize that it is possible to choose waveforms which are closer approximations to a sinusoidal waveform than is the square wave, but which are still suitable for this method. Fig. 2 shows a modulating

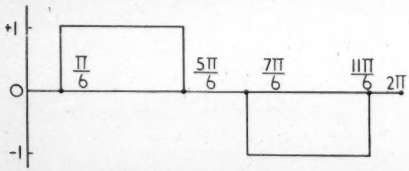


Fig. 2. Waveform with no second or third harmonics or their multiples

$$A_n = 2\sqrt{3}(a_n - a_{5n}/5 - a_{7n}/7 + a_{11n}/11 \ldots.)$$
$$n = 1$$

function which is almost as simple as Daniels' but which eliminates the contribution from the $3n^{\text{th}}$ harmonic and its multiples in the expression for A_n.

The only restriction on the waveform of $F(x)$ is that its steps must be integral, otherwise the planimeter operations would not be simple. Under this condition the waveform of $F(x)$ may be chosen to remove as many harmonics as possible from its Fourier expansion, thus reducing the number of corrections that must be applied to obtain a_n from A_n.

In the following argument the case $n = 1$ is considered, corresponding to one period of $F_1(x)$ in the interval 0 to 2π, but $F_n(x)$ can easily be obtained by repeating the period n times in the interval instead of once.

If the coefficients of the sine terms are to be found, then the cosine terms in the series for $F_1(x)$ have to be removed by making $F_1(x)$ anti-symmetric about π.

In general there are two ways of removing any given harmonic. First, $F_1(x)$ may be made symmetrical about $N\pi/k$ where k is the order of the harmonic and N is any integer. In this case all the multiples of the k^{th} harmonic will also be removed. For example, in both the simple square waveform of Fig. 1 and the waveform of Fig. 2, the second harmonic and its multiples, that is, all even harmonics, are removed by making the waveform symmetrical about $\pi/2$.

Secondly, the modulating waveform $F_1(x)$ may consist of pairs of identical elements separated by half the wavelength of the harmonic to be removed. This process will also remove the odd multiples of the harmonic. For example, the waveform in Fig. 2 consists of pairs of elements separated by $\pi/3$, that is, each element is half of one of the "blocks," and the third harmonic and its multiples are removed.

Thus, provided all the even harmonics are removed by the application of the first process, the application of the second process to a harmonic of order k will also remove all its multiples. It is worth while to note that after the last subdivision of the modulating waveform into pairs of elements the width of each element may be adjusted either to simplify the shape of $F_1(x)$ or to eliminate one more harmonic and its multiples by making the width of each element equal to a wavelength of the harmonic. Fig. 3(*a*) shows a waveform in which the first process has been used to eliminate all the

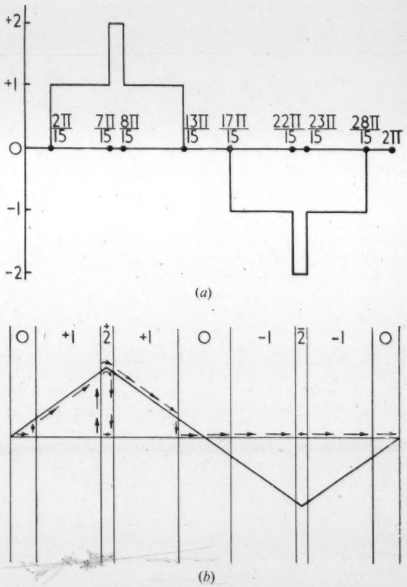


Fig. 3. (*a*) Waveform with no second, third or fifth harmonics or their multiples, and (*b*) the planimeter operation to measure A_1 using a triangular wave as an example (only outward path shown)

$$A_n = 8(\cos\pi/6)(\sin\pi/5)[a_n - 2(\cos\pi/5)a_{7n}/7 - a_{11n}/11$$
$$n = 1 \qquad + 2(\cos\pi/5)a_{13n}/13 \ldots]$$

even harmonics, the second to remove the third harmonic and its multiplies, and the width of the elements then adjusted to remove the fifth harmonic and its multiples.

In practice it will probably be better to minimize the effect of higher-order harmonics as a whole instead of eliminating individual harmonics completely. The waveform of Fig. 4 has been drawn with this in mind. The criterion assumed to give the total harmonic content of the waveform is the ratio of the square root of the sum of squares of the coefficients of harmonics above the first to the coefficient of the first harmonic. In the particular case of Fig. 4 the modulating function has been made symmetrical about $N\pi/3$ and of unit

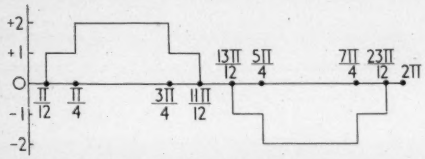


Fig. 4. Waveform with no third harmonic and small total harmonic content

$$A_n = 8(\cos \pi/12)(\cos \pi/6)/\pi \sum_{n=1} [a_n - (\tan \pi/12)a_{5n}/5 + (\tan \pi/12)a_{7n}/7 - a_{11n}/11 \ldots .]$$

magnitude. The width of each step has then been adjusted to make the above ratio a minimum. (The desired width is such that there is some overlapping of individual steps.) It is easily found that in the case of Fig. 4 the harmonic content is under 17%, whilst for the square wave (Fig. 1) it is 48%, and in Fig. 2 it is 31%. It is obvious then that the waveform in Fig. 4 will result in a considerable saving in time and is probably the most useful.

As successive harmonics are eliminated, the modulating waveform gets more complicated owing both to overlapping of the individual elements and to the fact that each set of harmonics removed is prime to those already eliminated. This implies that the sub-division involves fractions of π with an increasing product of primes in the denominator. In practice the optimum degree of complexity will be governed by the nature of the problem, and the labour of preparing the more complex grids will probably only be worth while when many curves with the same wavelength are to be analysed. There is no reason why all the grids should have the same complexity: the higher-order grids might well be simpler than those of lower order, though a factor would be necessary to make the conversion from measured areas into harmonic amplitude comparable.

The planimeter procedure is as follows. When the modulating function $F_n(x)$ is zero the planimeter follows the axis. When it is positive the instrument follows the curve in a positive sense, and when it is negative the curve is followed in a negative sense. This can all be done in one operation, including the allowance for double amplitude of $F_n(x)$, by following the curve as shown in Fig. 3(*b*) (only the outward path is shown). Here the analysis of a triangular wave is illustrated, and this demonstrates the advantage of the improved method. Using Daniels's method, the value of A_1 obtained contains approximately a 5% contribution from the higher harmonics. Using the modulating function of Fig. 2, the value of A_1 contains approximately a 1% contribution from higher harmonics, and using the modulating function of Fig. 4, the contribution from the higher harmonics is approximately 0·2%.

It can be seen that the gain using the modulating function of Fig. 4, compared with the square modulating function, is much greater than might be expected from the values for the harmonic content of these functions. This is due to the fact that corrections are moved to the higher and less important harmonics. As before, the cosine terms may be found by shifting $F_n(x)$ to the right by $\pi/2n$.

CONCLUSION

Daniels's method of harmonic analysis has been analysed in a way that leads to further development. The improved methods suggested greatly reduce the number of corrections which have to be applied to obtain the amplitude of a harmonic component from an area measured by the planimeter, and, in many practical cases in which the amplitudes of the harmonics decrease rapidly with increasing order a single planimeter operation, will give a sufficiently accurate value of a harmonic amplitude without correction. In cases where many curves are to be analysed a considerable saving in time will result.

APPENDIX

Worked example

In general, four independent series of results have to be worked out to give (1) sine components of odd harmonics, (2) sine components of even harmonics, (3) cosine components of odd harmonics and (4) cosine components of even harmonics. To demonstrate the procedure a saw-tooth waveform, which is comparatively rich in harmonics, will be analysed. It will be assumed that the amplitudes of the first five harmonics are to be measured to an accuracy of 1% of the fundamental. The waveform analysed has a length of 36 cm and an amplitude of 10 cm (Fig. 5). Here, we shall only deal with the first series of results.

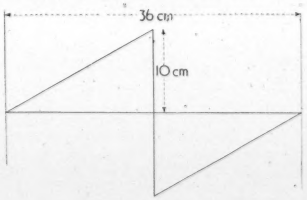


Fig. 5. Waveform analysed in the Appendix

By inspection of the waveform it is usually possible to gain some idea of the amplitude of the harmonics, and in this case it would be safe to assume that the amplitude decreased steadily with increasing order. We shall therefore start by ignoring coefficients above the eleventh when correcting the measured areas.

The modulating function used is that of Fig. 4 and grids are prepared accordingly. They are placed over the waveform to be analysed and the areas measured with a planimeter as described above. The areas are:

$$A_1 = 241{\cdot}4\ \text{cm}^2 \qquad A_3 = 79{\cdot}2\ \text{cm}^2 \qquad A_5 = 48{\cdot}8\ \text{cm}^2$$

The uncorrected coefficients (the first term in the series for A_n given under Fig. 4) obtained from these areas are:

$$a_1 = 6{\cdot}30\ \text{cm} \qquad a_3 = 2{\cdot}07\ \text{cm} \qquad a_5 = 1{\cdot}27\ \text{cm}$$

The first correction for a_3 will be approximately 1/20 of a_{15} and we may assume from an inspection of a_1, a_3, a_5 that this is negligible to the limits of accuracy that have been set. The same argument applies to a_5. Only a_1 needs to be corrected. As the correction due to a_{11} is $a_{11}/11$, it will probably be necessary to calculate both a_7 and a_{11}. (a_5 is already known.) Only rough values are required as they are being used for small corrections so we may use the simpler square modulating function to calculate them.

The areas are:

$$A_7 = 27\ \text{cm}^2 \qquad A_{11} = 17\ \text{cm}^2$$

and it follows from the expression in the legend of Fig. 1(*b*) that

$$a_7 = 1{\cdot}18\ \text{cm} \qquad a_{11} = 0{\cdot}73\ \text{cm}$$

Thus

$$a_1 = 6{\cdot}30 + 0{\cdot}054a_5 - 0{\cdot}038a_7 + 0{\cdot}091a_{11} = 6{\cdot}39\ \text{cm}$$

To be sure of 1% accuracy we should, strictly, work out the next correction in each case. However, the method of working the measurements has been demonstrated, so this will not be done.

The complete results are now given with the theoretical amplitudes in brackets for comparison, and they give some idea of the accuracy of the method.

$$\begin{aligned} a_1 &= 6{\cdot}39\ \text{cm}\ (6{\cdot}37) \\ a_3 &= 2{\cdot}07\ \text{cm}\ (2{\cdot}12) \\ a_5 &= 1{\cdot}27\ \text{cm}\ (1{\cdot}27) \\ a_7 &= 1{\cdot}18\ \text{cm}\ (0{\cdot}91) \\ a_{11} &= 0{\cdot}73\ \text{cm}\ (0{\cdot}58) \end{aligned}$$

a_7 and a_{11} are in error owing to the use of only a crude approximation.

An application of conformal transformation to the investigation of the magnetic field between galvanometer pole-pieces

By W. Snowdon, M.Sc., and N. Davy, D.Sc., University of Nottingham

[*Paper first received* 5 *August and in final form* 22 *September*, 1953]

An accurate calculation is made of the magnetic potential, the associated stream function, and the strength of magnetic field along the boundary and axes of symmetry of a figure which represents the cross-section of the two pole-pieces and central core of a galvanometer. In addition the field strength along a locus in the space between the core and pole-pieces is determined precisely. Part of this locus lies along the path of the moving coil of the galvanometer. The method used is an extension of the usual procedure of combining Schwarz–Christoffel transformations. The paper includes a map of the field obtained by an experimental method which employs the calculated values of the stream-potential function round the boundary of the figure.

A conformal transformation expressed by means of an analytic function of a complex variable defines a correspondence between orthogonal systems of curves in one plane and perpendicular systems of straight lines in another, and it has been pointed out by Richmond[1] that this fact may be used to convert a two dimensional field mapping problem involving conductors or pole-pieces with curved boundaries into one involving straight boundaries. In the paper referred to, Richmond has applied this to cylindrical conductors having cross-sections including arcs of conic sections. The same principle has been applied by Adams[2,3,4] to problems concerning conductors bounded by parts of circular cylinders.

The object of this paper is to use Richmond's method to investigate the field of two cylindrical pole-pieces with a central circular cylindrical core as they are usually arranged in a galvanometer.

THE TRANSFORMATION FROM CURVED TO STRAIGHT BOUNDARIES

The transformation made use of here is

$$\zeta = z - 1/z \tag{1}$$

Corresponding to the straight line $\xi = a$ in the ζ-plane, there is a curve

$$x(x^2 + y^2) - a(x^2 + y^2) - x = 0 \tag{2}$$

in the z-plane. This is symmetrical about the x-axis and consists of two parts. The line $x = a$ is an inflexional asymptote to one part, which is confined between $x = a$ and $x = \frac{1}{2}[a + \sqrt{(a^2 + 4)}]$. The other part is a closed curve between $x = 0$ and $x = \frac{1}{2}[a - \sqrt{(a^2 + 4)}]$. Each part is the reflexion in the y-axis of the inverse of the other with respect to the circle $|z| = 1$. The curve $\xi = 0$ in particular, consists of the y-axis and the circle $|z| = 1$.

Corresponding to the straight line $\eta = b$ in the ζ-plane, there is in the z-plane a curve

$$y(x^2 + y^2) - b(x^2 + y^2) + y = 0 \tag{3}$$

This is symmetrical about the y-axis and has three forms depending on the value of b. In each of its three forms the parts of the curves inside and outside the circle $|z| = 1$ are inverse with respect to this circle, and the line $y = b$ is an inflexional asymptote. When b is less than 2, there is one continuous curve between its asymptote $y = b$ and the x-axis, and it is this form which is relevant to the present investigation.

It will be noted that outside the circle $|z| = 1$ the lines $\xi =$ constant are lines of flow round a circular cylindrical obstacle of otherwise uniform flow parallel to the y-axis. It was this hydrodynamic analogy which led to the choice of the particular transformation used here, by reason of its obvious connexion with the geometry of the pole-piece surfaces and central core of a galvanometer. From the pattern of ortho-

gonal sets of curves in the z-plane produced by giving a and b different values in equations (2) and (3) above, it is possible to select parts of curves to build up the outline required. These component parts are assembled in Fig. 1 and enumerated below.

(*a*) The circle $|z| = 1$ from the locus $\xi = 0$.

(*b*) The portions of $\xi = \pm a$ (where a is about 0·5) outside the circle $|z| = 1$ between $\eta = \pm b$ (where b is about 1·5). These are *EDG* and *MN* in Fig. 1.

(*c*) The portions of $\eta = \pm b$ remote from the origin cut off by $\xi = \pm a$. These are *LM*, *EF*, *PN*, and *GH* in Fig. 1.

In effect, the transformation

$$\zeta = z - 1/z$$

has been employed to set up a correspondence between the set of orthogonal curves in Fig. 1 and the set of perpendicular

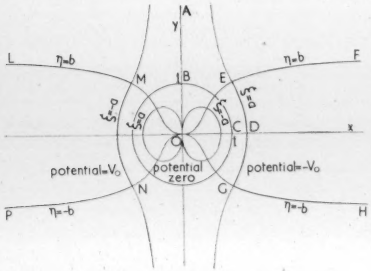


Fig. 1. Cross-section in z-plane of pole-pieces and core

straight lines in Fig. 2. It is sufficient in practice to consider the region to the left of the boundary *ABCDEF* in both figures.

At this stage, it is desirable to outline briefly the subsequent procedure. Schwarz–Christoffel transformations are derived which relate (*a*) the region *ABCDEF* in the ζ-plane to the upper half of the t-plane (the geometrical transformation), and (*b*) the upper half of the t-plane to the appropriate region of the W-plane, where $W = U + iV$ is the stream-

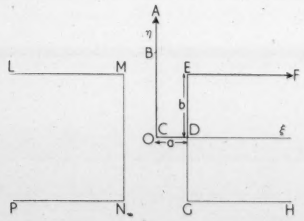


Fig. 2. The transformed figure in the ζ-plane

potential function (the magnetic transformation). From these, expressions are obtained for the co-ordinates of points in the z-plane and the values of the stream-potential function and field strength at these points, in terms of elliptic functions of a parameter $w = u + iv$, connected with t by the equation $t = sn^2 w$. The use of the intermediate variable t has already been described by Snowdon and Davy[5] and need receive no further mention here.

THE GENERAL FORMULAE

In the paper [ref. (5)], it has been shown that the transformation

$$\zeta = (a/E)E(w) \tag{4}$$

relates conformally the domain *ABCDEF* in the ζ-plane with the interior and boundary of the rectangle *ABCDEF* in

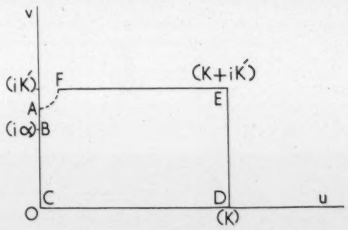


Fig. 3. The w-plane showing boundary values of the parameter w

the w-plane shown in Fig. 3. The modulus of the elliptic function is given by

$$b/a = (K' - E')/E \tag{5}$$

Suppose the potential difference between the central core and one pole-piece in Fig. 1 is V_0. The domain *ABCDEF* in the w-plane is transformed into the interior and boundary of the

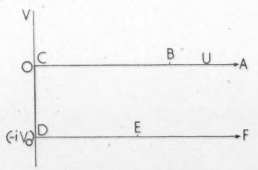


Fig. 4. The W-plane showing boundary values of the stream-potential function W

semi-infinite strip *ABCDEF* in the W-plane in Fig. 4 by the relation

$$dW/dw = -(2V_0 i\pi)\, dnw \tag{6}$$

In the z-plane therefore, the field strength at any point is

$$|dW/dz| = |(dW/dw)(dw/d\zeta)(d\zeta/dz)|$$

$$\text{i.e. field strength} = (2V_0 E/\pi a)\sqrt{(\zeta^2 + 4)/z\, dn\, w} \tag{7}$$

Finally, integrating equation (6) with the aid of equation (4), taking W zero at the point C and $-iV_0$ at the point D, the following general formula for the stream-potential function is obtained.

$$W = -i(2V_0/\pi)\sin^{-1}(sn\, w) \tag{8}$$

THE FORMULAE ALONG PARTICULAR LOCI

(a) *Along AB*

The variable z is purely imaginary along AB, and equal to iy, where

$$y = \tfrac{1}{2}[\eta + \sqrt{(\eta^2 - 4)}] \tag{9}$$

The value of ζ is also purely imaginary, being $i\eta$, where

$$\eta = (a/E)[\mathrm{sn}'v\mathrm{dn}'v/\mathrm{cn}'v + v(K' - E')/K' - Z'(v)] \tag{10}$$

dashes denoting functions to the complementary modulus. The variable v takes real values between K' (at A) and a value α (at B) intermediate between K' and zero, such that

$$\mathrm{sn}'\alpha\mathrm{dn}'\alpha/\mathrm{cn}'\alpha + \alpha(K' - E')/K' - Z'(\alpha) = 2E/a \tag{11}$$

The most convenient form of equation (7) for the field strength along AB is

$$F = (2V_0E/\pi a)\sqrt{(\eta^2 - 4)}\,\mathrm{cn}'v/\mathrm{dn}'v \tag{12}$$

and the stream function may be expressed as

$$U = (2V_0/\pi)\cosh^{-1}(\mathrm{nc}'v) \tag{13}$$

The field strength at B is clearly zero.

(b) *Along BC*

The angle made by a radius of the quadrant BC with the x-axis may be shown to be θ, where

$$\sin\theta = \eta/2 \tag{14}$$

η being given by equation (10), in which v takes real values between α (at B) and zero (at C). Further, since $|z| = 1$ along the quadrant,

$$\text{field strength} = (2V_0E/\pi a)\sqrt{(4 - \eta^2)}\,\mathrm{cn}'v/\mathrm{dn}'v \tag{15}$$

and the stream function is expressed by equation (13) as before.

(c) *Along CD*

Along the x-axis, z may be replaced by x, so that

$$x = \tfrac{1}{2}[\xi + \sqrt{(\xi^2 + 4)}] \text{ where } \xi = (a/E)E(u) \tag{16}$$

the real variable u increasing from zero (at C) to K (at D).

$$\text{Also, field strength} = (2V_0E/\pi a)\sqrt{(\xi^2 + 4)}/x\,\mathrm{dn}\,u \tag{17}$$

$$\text{and potential} = V = -(2V_0/\pi)\sin^{-1}(\mathrm{sn}u) \tag{18}$$

The expressions for the field strength at C and D, obtainable from equation (17), are

$$\text{field strength at } C = 4V_0E/\pi a \tag{19}$$

$$\text{field strength at } D = (V_0E/\pi ak')\ [a^2 + 4 - a\sqrt{(a^2 + 4)}] \tag{20}$$

(d) *Along DE*

On the inner surface DE of the pole-piece, the parameter w is equal to $K + iv$, the value of v varying from zero to K'. Thus $\xi = a$, and

$$\eta = (a/E)\{[\mathrm{sn}'v\mathrm{dn}'v/\mathrm{cn}'v + v(K' - E')/K' - Z'(v)] - k^2\,\mathrm{sn}'v/\mathrm{cn}'v\mathrm{dn}'v\} \tag{21}$$

The position of the corresponding point in the z-plane is more advantageously expressed in polar co-ordinates (r, θ), which, by solving simultaneously the polar forms of equations (2) and (3), may be proved to be given by

$$\cos 2\theta = \tfrac{1}{4}\{\sqrt{[(\xi^2 + \eta^2)^2 + 8(\xi^2 - \eta^2) + 16]} - (\xi^2 + \eta^2)\} \tag{22}$$

$$\text{and} \qquad r = \tfrac{1}{2}(\xi\sec\theta + \eta\,\mathrm{cosec}\,\theta) \tag{23}$$

Associated with each point thus determined, we have

$$\text{field strength} = (2V_0E/\pi ak')[(a^2 + \eta^2)^2 + 8(a^2 - \eta^2) + 16]^{\frac{1}{4}}\mathrm{dn}'v/r\mathrm{cn}'v \tag{24}$$

$$\text{stream function} = U = (2V_0/\pi)\cosh^{-1}(\mathrm{nd}'v) \tag{25}$$

At the point E, $\xi = a$, $\eta = b$, the field is of infinite strength, and the stream function is

$$U = (2V_0/\pi)\cosh^{-1}(1/k) \tag{26}$$

(e) *Along EF*

The formulae for points on the upper surface EF of the pole-piece result from the substitution $w = u + iK'$, in which u starts with the value K at E and decreases as the point moves towards the right. $\eta = b$ and

$$\xi = (a/E)[E(u) + \mathrm{cn}u\mathrm{dn}u/\mathrm{sn}u] \tag{27}$$

Equations (22) and (23) determine the polar co-ordinates of the points in the z-plane. Also

$$\text{field strength} = (2V_0E/\pi a)[(\xi^2 + b^2)^2 + 8(\xi^2 - b^2) + 16]^{\frac{1}{4}}\mathrm{sn}u/r\mathrm{cn}u \tag{28}$$

and

$$\text{stream function} = U = (2V_0/\pi)\cosh^{-1}(1/k\mathrm{sn}u) \tag{29}$$

(f) *Along the locus* $u = K/2$

Before giving the formulae associated with points for which $u = K/2$, some explanation of the reason for choosing this particular locus would appear to be called for. These points lie on a curve which, starting from near the mid-point of CD, assumes almost exactly the shape of a circle concentric with the central core for an angular distance of about 30°. Thereafter it moves round the point E, terminating on the upper surface of the pole-piece very close to E (see Fig. 5). It therefore supplies the means of calculating the field strength along the path of a moving coil suspended between the pole-pieces for a considerable angle on each side of its mean position.

The formulae obtained by substituting $w = K/2 + iv$ are unwieldy, and in order to state them more conveniently, the following definitions are made.

$$S = k^2/(1 + k') \tag{30}$$

$$Q = (\mathrm{cn}'v/\mathrm{sn}'v)^2 + S \tag{31}$$

$$R^2 = (\xi^2 + \eta^2)^2 + 8(\xi^2 - \eta^2) + 16 \tag{32}$$

Using these definitions, the expressions become

$$\xi = a/2 + (aS/E)(\tfrac{1}{2} + k'/Q) \tag{33}$$

$$\eta = (a/E)\{[\mathrm{sn}'v\mathrm{dn}'v/\mathrm{cn}'v + v(K' - E')/K' - Z'(v)] - S\mathrm{dn}'v/Q\mathrm{sn}'v\mathrm{cn}'v\} \tag{34}$$

and field strength

$$= [2V_0E/\pi a\sqrt{(k')}]Q(\mathrm{sn}'v)^2\sqrt{R}/r\sqrt{[(\mathrm{cn}'v\mathrm{dn}'v)^2 + S^2(\mathrm{sn}'v)^2]} \tag{35}$$

The transformation from the ζ-plane to the z-plane is effected as before by means of the polar equations (22) and (23).

At the point $w = K/2$, where the above locus meets the x-axis, the value of ξ is $\frac{1}{2}a(1 + S/E)$ and the field strength and potential may be calculated from equations (16), (17) and (18), in which u is replaced by $K/2$. Similarly at the end of the curve on EF, by the substitution $w = K/2 + iK'$, $\xi = \frac{1}{2}a(1 + S/E + 2k'/E)$, and the other quantities are most easily obtained by putting $u = K/2$ in equations (28) and (29).

THE CALCULATIONS

In order that the profile of pole-pieces and central core should approximate as closely as possible to that occurring in practice, the ratio b/a should be about 3. In this case, k is approximately 0·0125. For this value of modulus, approximations to the elliptic functions in terms of circular and hyperbolic functions are available, which are adequate for use with six-figure tables. The formulae required, with two exceptions, may be found in a recent work[6], the two additional formulae needed being

$$E(u,k) = u - (k^2/4)(2u - \sin 2u) \qquad (36)$$

$$E(v, k') = \tanh v + (k^2/4)(v - \tanh v + v\tanh^2 v) \qquad (37)$$

The approximation consists in ignoring terms containing the fourth and higher powers of k. By the use of these formulae the values which may be assigned to a and b are theoretically open to unrestricted choice, although the calculations for any points not on the boundary $ABCDEF$ are extremely arduous. In the present work, the values sin 1° and sin 89° have been chosen for k and k' respectively. This choice has permitted the use of the Smithsonian tables for elliptic functions having the modulus sin 89°, although the use of the approximations described above for elliptic functions of modulus sin 1° was unavoidable. The actual profile for these values and the locus $u = K/2$ are drawn accurately in Fig. 5.

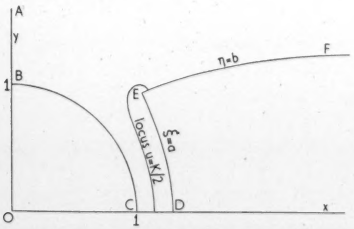


Fig. 5. Positive quadrant of z-plane for $a = 0{\cdot}5$, $b = 1{\cdot}4115$

The polar co-ordinates of points along the boundary $ABCDEF$ have been calculated by giving various values to the parameter w. At each point the field strength and potential or stream function have been determined. The polar co-ordinates of points on the locus $u = K/2$ and the field strength at each point have found. Field strengths and stream-potential functions are expressed as multiples of V_0, the potential difference between the pole-piece and core. The unit of length is the radius of the core. To obtain the results for a core of radius ρ units, all lengths are to be multiplied by ρ, field strengths given here are to be divided by ρ, while the values of the stream-potential function are unaltered. The results are set out in Tables 1 to 6, and illustrated graphically in Figs. 6, 7 and 8.

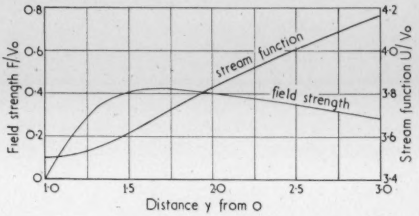


Fig. 6. Graphs of field strength and stream function along AB (on y-axis)

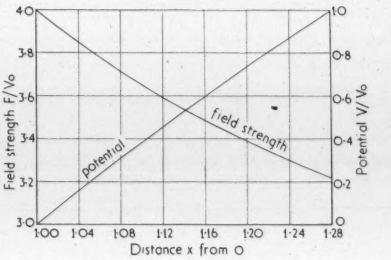


Fig. 7. Graphs of field strength and potential along CD (on x-axis)

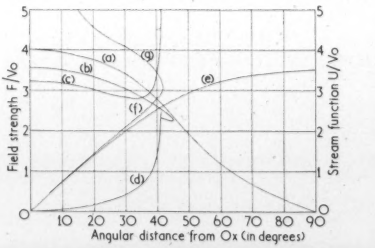


Fig. 8. Graphs of field strength and stream function along surfaces of pole-pieces and locus $u = K/2$

Field strength along
- (*a*) surface of core (quadrant *BC*).
- (*b*) curve $u = K/2$.
- (*c*) inner surface of pole-piece (*DE*).
- (*d*) upper surface of pole-piece (*EF*).

Stream function along
- (*e*) surface of core (*BC*).
- (*f*) inner surface of pole-piece (*DE*).
- (*g*) upper surface of pole-piece (*EF*).

Table 1. *Field strength* F *and stream function* U *at points* (O, y) *on AB*

$90v/K'$	90	87	86	85	84	83	*at B*
y	∞	2·7691	2·2295	1·8655	1·5800	1·3125	1
F/V_0	0	0·3116	0·3786	0·4177	0·4163	0·3349	0
U/V_0	∞	4·1026	3·9167	3·7712	3·6509	3·5478	–

Table 2. *Field strength* F *and stream function* U *at points having polar co-ordinates* $(1,\theta)$ *on the quadrant BC*

$90r/K'$	82	81	80	78	75
θ	79°36′	70° 7′	64°46′	57°55′	51°32′
F/V_0	0·3238	0·6740	0·9207	1·3168	1·7893
U/V_0	3·4572	3·3756	3·3014	3·1688	2·9967
$90r/K'$	70	65	60	50	40
θ	44°45′	39°52′	35°50′	28°53′	22°39′
F/V_0	2·3750	2·7842	3·0744	3·4469	3·6738
U/V_0	2·7505	2·5306	2·3237	1·9272	1·5491
$90r/K'$	30	20	10	0 (*at C*)	
θ	16°46′	11°5′	5°31′	0	
F/V_0	3·8244	3·9237	3·9809	3·9997	
U/V_0	1·1537	0·7690	0·3845	0	

Table 3. *Field strength* F *and potential* V *at points* (x,0) *on CD*

u	$\pi/12$	$\pi/6$	$K/2$	$\pi/3$	$5\pi/12$	K (*at D*)
x	1·0425	1·0868	1·1328	1·1805	1·2298	1·2808
F/V_0	3·8399	3·6931	3·5586	3·4354	3·3226	3·2195
$-V/V_0$	0·1667	0·3333	0·5000	0·6664	0·8333	1

Table 4. *Field strength* F *and stream function* U *at points* (r,θ) *on surface DE*

$90r/K'$	10	20	30	40	50
θ	5°21′	10°44′	16°11′	21°45′	27°24′
r	1·2821	1·2863	1·2937	1·3047	1·3205
F/V_0	3·2039	3·1574	3·0810	2·9789	2·8670
U/V_0	0·3844	0·7687	1·1528	1·5361	1·9170
$90r/K'$	60	65	70	80	90 (*at E*)
θ	33° 0′	35°39′	38° 3′	41°18′	41°59′
r	1·3416	1·3539	1·3666	1·3867	1·3914
F/V_0	2·7994	2·8305	2·9728	4·3911	∞
U/V_0	2·2897	2·4683	2·6364	2·9088	3·0184

Table 5. *Field strength* F *and stream function* U *at points* (r,θ) *on surface EF*

u	$K/2$	$\pi/6$	$\pi/12$	$\pi/18$
θ	41°19′	39°45′	34°3′	28°55′
r	1·4475	1·5706	2·0277	2·5225
F/V_0	2·3242	1·3225	0·5968	0·3850
U/V_0	3·2391	3·4598	3·8790	4·1330

Table 6. *Field strength* F *at points* (r,θ) *on the locus* $u = K/2$

$90r/K'$	0	10	20	30	40
θ	0	5°28′	11°0′	16°37′	22°23′
r	1·1328	1·1335	1·1355	1·1392	1·1452
F/V_0	3·5586	3·5408	3·4871	3·3959	3·2642
$90r/K'$	50	60	70	75	78
θ	28°24′	34°41′	41°0′	43°49′	44°57′
r	1·1554	1·1754	1·2234	1·2726	1·3142
F/V_0	3·0865	2·8526	2·5554	2·3937	2·3129
$90r/K'$	80	82	85	87	90
θ	45°19′	45°17′	44°22′	43°15′	41°19′
r	1·3466	1·3804	1·4258	1·4451	1·4475
F/V_0	2·2733	2·2496	2·2485	2·2695	2·3242

From the graphs of the stream-potential function, it is possible to determine the position on the boundary *ABCDEF* of the ends of equipotential lines and lines of force at stated intervals in the values of U and V. This facilitates greatly the experimental determination of a map of the field using the electrodynamic analogy. Fig. 9 was obtained by this method using heavy brass electrodes on thick soaked blotting

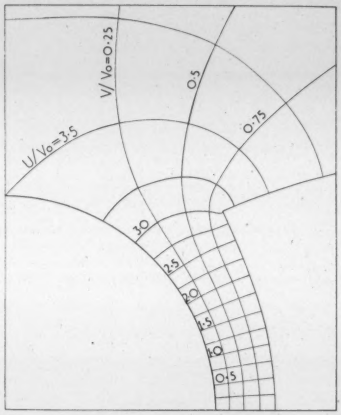


Fig. 9. Map of the field

paper, and confirms in a general way the two main features of the variation of the field in the gap between pole-piece and central core exhibited by curves (*a*), (*b*) and (*c*) in Fig. 8. These features are first, the decrease in field strength as the gap increases, and second, the slight fall in field strength across the gap along a radius vector. The method used to obtain the plot does not admit of a quantitative check of the second of these, but an approximate comparison of theory and experiment has been made in the case of the first feature. This was made possible by the fact that up to the line of force $V = 2{\cdot}5\ V_0$ the equipotential line $U = 0{\cdot}5\ V_0$ and the locus $u = K/2$ are nearly coincident. The field strengths along $U = 0{\cdot}5\ V_0$ at each intersection with a line of force have been estimated by taking the quotient of V_0 and twice the distance along the line of force between the experimentally obtained equipotential lines $U = 0{\cdot}25\ V_0$ and $U = 0{\cdot}75\ V_0$.

Table 7. *Field strength halfway between pole-piece and central core as obtained by three methods*

(*a*) Values obtained from experimental plot.
(*b*) Theoretical values along locus $u = K/2$ by interpolation from curve (*b*) in Fig. 8.
(*c*) Values calculated from formula (38) on the assumption that the gap may be treated as a circular annulus of radii equal to the radii vectors of the gap.

Co-ords. of point					
	r	1·134	1·136	1·138	1·140	1·142
	θ	3½°	7°	11°	14½°	18°
(*a*)		$3{\cdot}57V_0$	$3{\cdot}57V_0$	$3{\cdot}52V_0$	$3{\cdot}47V_0$	$3{\cdot}42V_0$
(*b*)		$3{\cdot}55V_0$	$3{\cdot}53V_0$	$3{\cdot}49V_0$	$3{\cdot}44V_0$	$3{\cdot}37V_0$
(*c*)		$3{\cdot}57V_0$	$3{\cdot}57V_0$	$3{\cdot}51V_0$	$3{\cdot}46V_0$	$3{\cdot}40V_0$
Co-ords. of point	r	1·144	1·150	1·156	1·164	1·180
	θ	22°	25½°	29½°	34°	38°
(*a*)		$3{\cdot}33V_0$	$3{\cdot}21V_0$	$3{\cdot}13V_0$	$2{\cdot}98V_0$	$2{\cdot}78V_0$
(*b*)		$3{\cdot}28V_0$	$3{\cdot}18V_0$	$3{\cdot}04V_0$	$2{\cdot}88V_0$	$2{\cdot}73V_0$
(*c*)		$3{\cdot}33V_0$	$3{\cdot}24V_0$	$3{\cdot}10V_0$	$2{\cdot}95V_0$	$2{\cdot}76V_0$

These results are compared in Table 7 with the values of field strength at the same points (or points very close to these) read by interpolation from the theoretical curve (*b*) in Fig. 8. There is no significant difference between the experimental and theoretical values.

CONCLUSIONS

In the interpretation of the above results, it should be noted that although the inner surface of the pole-piece *DE* and the locus $u = K/2$ are not exactly circular, their departure from a circle is at first only slight. In fact, at an angular distance of $27\frac{1}{2}°$ from the *x*-axis the radius vector to the inner surface *DE* has increased by 3%, and the radius to the locus $u = K/2$ by 2%. If *DE* were the arc of a circle and *EF* a straight line parallel to the *x*-axis, the problem could not be treated in the manner above because these curves are no longer orthogonal. In this case it is not possible to produce by known methods an exact solution in terms of standard mathematical functions. However, from the results obtained here, the following general assertions can be made which apply to the field ideally produced by the ordinary arrangement of magnets in a galvanometer.

Evidently the influence of the vertex at *E* in raising the field strength is only great for points very close to *E*; the field strength at any point on the curve encircling the pole *E* shown in Fig. 5 is less than that at *C*, and actually remains almost constant as the pole is approached round the curve; and the field strength along the inner surface of the pole-piece is less than that at *C* to within one degree of the vertex *E*. It is of practical interest therefore to compare the actual field with that calculated on the assumption that the gap between pole and central core along any radius vector may be treated as lying between concentric circles; i.e. that

$$\text{field strength at } P = (V_0/r)\log(r_2/r_1) \qquad (38)$$

where r, r_2, r_1 are the radii to the point *P*, the surface of the pole-piece, and the surface of the central core respectively. A simple calculation shows that field strengths obtained from the above formula (38) are:

(i) increasingly smaller than the true theoretical values as the point *P* moves along the inner surface of the pole-piece away from the *x*-axis, e.g. $2{\cdot}71\ V_0$ instead of $2{\cdot}82\ V_0$ at $(1{\cdot}33, 29\frac{1}{2}°)$,

(ii) increasingly higher than the true values as the point *P* moves along the surface of the core away from the *x*-axis, e.g. $3{\cdot}58\ V_0$ instead of $3{\cdot}43\ V_0$ at $(1, 29\frac{1}{2}°)$,

(iii) nearly identical with the true values at points halfway across the gap up to an angular distance of 38° from the *x*-axis. This last deduction is of practical importance and the extent of agreement is shown in Table 7, where field strengths along the equipotential line $U = 0{\cdot}5\ V_0$ obtained from the elementary formula (38) are given for comparison alongside the true values and the experimental values already mentioned above.

REFERENCES

(1) RICHMOND, H. W. *Proc. Lond. Math. Soc.*, 2, **22**, pp. 483–494 (1924).

(2) ADAMS, E. P. *Proc. Amer. Phil. Soc.*, **75**, pp. 11–70 and 549–57 (1935).

(3) ADAMS, E. P. *Proc. Amer. Phil. Soc.*, **76**, pp. 251–302 (1936).

(4) ADAMS, E. P. *Proc. Amer. Phil. Soc.*, **78**, pp. 191–243 and 245–70 (1937).

(5) SNOWDON, W., and DAVY, N. *Brit. J. Appl. Phys.* **4**, pp. 339–41 (1953).

(6) MILNE-THOMSON, L. *Jacobian Elliptic Function Tables.* (New York: Dover Publications Inc., pp. 20–21, 1950).

A magnetic method for the determination of γ-loops in binary iron alloys and its application to the iron-silicon system

By J. CRANGLE, Ph.D., A.Inst.P., Physics Department, University of Sheffield

[*Paper first received* 14 *August, and in final form* 11 *September*, 1953]

A new method of phase analysis which depends on measurements of paramagnetic susceptibility at different temperatures above the Curie point has been evolved. It is capable of indicating equilibrium conditions in two-phase alloys. The application to the iron-silicon system shows that here the γ-loop extends to a maximum of only 2·15% of silicon, and not 2·5% as previously reported.

THE APPLICATION OF PARAMAGNETIC SUSCEPTIBILITY MEASUREMENTS TO PHASE ANALYSIS

In pure iron at temperatures above the Curie point the α-phase (body-centred cubic) and the γ-phase (face-centred cubic) are both paramagnetic, but their susceptibilities are widely different. This difference is often maintained when other elements are dissolved in iron in binary alloys, and it forms the basis for a new method of determining the boundaries of the γ-loop regions in the equilibrium diagrams of such alloy systems.

In two-phase alloys where both phases are paramagnetic, the susceptibility of each phase acts independently. At a given temperature the mass susceptibility χ of a mixture of the γ- and α-phases is given by $\chi = y\chi_\gamma + (1 - y)\chi_\alpha$, where y is the proportion by weight of γ-phase present, and χ_γ and χ_α refer to the γ- and the α-phases respectively at the temperature chosen. χ_γ and χ_α are not independent of temperature, but they vary smoothly in a way which allows extrapolation from single-phase conditions to temperatures where mixed phases exist. Thus a measure of the proportions of the two phases present in a given state can be obtained.

Usually the α- to γ-phase change shows marked temperature hysteresis. The course of the graph showing the proportion of γ-phase as a function of temperature for this transformation would depend on whether the susceptibility was measured during heating or during cooling, and also on the rate of change of temperature. The graph for equilibrium conditions, corresponding to an infinitely slow rate, would lie somewhere between those for heating and cooling. The following method allows equilibrium conditions to be recognized.

If a specimen is heated from a low temperature into the region of the phase transformation and then the heating process is halted, the specimen will be in a non-equilibrium two-phase state possessing less γ-phase than the equilibrium amount. If it is then cooled slowly at a constant rate, the

tendency will be for the amount of γ-phase to continue to increase at a diminishing rate until the temperature has been reached at which the proportions of phases present are in equilibrium, after which it will decrease. When the amounts of γ-phase present during this slow cooling are plotted on a graph against temperature, the maximum of the line obtained represents a point on the graph of the equilibrium proportions of phases against temperature. The procedure can be repeated for a series of different initial states, and the equilibrium graph for each alloy is given by the locus of the maxima of the curves obtained. A similar procedure may be applied starting from states containing more γ-phase than the equilibrium amount and slowly heating the specimen. This technique is illustrated in a later section by its application to the iron-silicon system.

Phase boundaries can also be located roughly by observing discontinuities in the variation of susceptibility with temperature.

PREVIOUS DETERMINATIONS OF THE IRON-SILICON γ-LOOP

The accepted equilibrium diagram of the γ-loop region of the iron-silicon system as given by Hanson[1] is based largely on that determined by Oberhoffer and Kreutzer[2] in 1928. Preliminary high-temperature susceptibility measurements on some commercial iron-silicon alloys in which very large single crystals had been grown by Martindale and Langford[3] showed that such materials do not agree with this diagram, and in view of the age of the previous determination it seemed desirable that its reliability should be investigated. The new method was therefore applied to this system, using materials of high purity.

Oberhoffer and Kreutzer[2] employed the X-ray disappearing phase method using a high-temperature camera, in which the specimen was heated in a high vacuum. The most useful of the other previous determinations of this part of the phase diagram seem to be those of Esser and Oberhoffer[4] (dilatometric methods), Ruer and Klesper[5] (thermal analysis) and Wever and Giani[6] (thermal analysis). None of these last three covered a sufficiently wide range of compositions to include the whole γ-loop and it is not clear whether their experimental points do indicate the actual position of the outer phase boundary, but the trend of all was towards a narrower loop than that found by Oberhoffer and Kreutzer.[2] Fig. 1 summarizes the results of these earlier determinations.

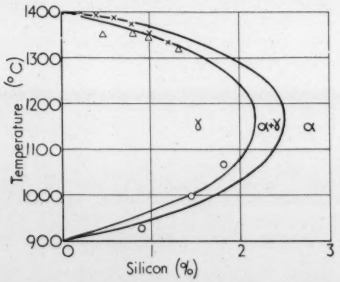


Fig. 1. Previous determinations of the γ-loop. The full lines are due to Oberhoffer and Kreutzer,[2] and the other points to Esser and Oberhoffer[4] (○), Ruer and Klesper[5] (×), and Wever and Giani[6] (△)

In the diagram published by Oberhoffer and Kreutzer, the apparent agreement between their work and that of Esser and Oberhoffer[4] appears to be the unfortunate consequence of an error in drawing.

MATERIALS USED

A series of iron-silicon alloys of very high purity, having silicon contents of up to 5% in steps of 1%, was obtained from the National Physical Laboratory. Their analyses were:

	34*AF*1 %	33*AF*3 %	33*AF*2 %	33*AF*1 %	33*AF*3 %
carbon	0·001 6	0·001 4	0·002 6	0·002 1	0·002 4
silicon	1·04	2·12	3·05	4·13	4·98
manganese	<0·005	<0·005	<0·005	<0·005	0·003
sulphur	0·005 2	0·004 9	0·004 9	0·004 5	0·006
phosphorus	—	—	—	0·002	—
nickel	—	—	—	0·005	—
chromium	—	—	—	0·001	—
copper	—	—	—	0·004	—
aluminium	0·001	0·002	0·001	<0·001	0·004
oxygen	0·001 1	0·000 94	0·001 0	0·000 87	0·001 1
nitrogen	0·001 3	0·001 2	0·001 4	0·001 2	0·000 6
hydrogen	—	—	—	<0·000 005	—

Of these, only the first two were capable of being taken into the γ-loop region. Three more alloys, with silicon contents of between 1% and 2% were made by melting parts of the higher members of the series with chosen proportions of deoxidized high-purity iron AH from the British Iron and Steel Research Association, for which the analysis was:

	%
carbon	0·002
manganese	<0·005
silicon	0·004
sulphur	0·004
phosphorus	<0·001
nickel	0·012
copper	0·005
aluminium	0·003
oxygen	0·001 6
nitrogen	0·001 4

Melting was carried out *in vacuo* in an induction furnace using crucibles of recrystallized alumina (Morgan "Triangle RR"). Each melt weighed about 40 g.

All three alloys were subsequently homogenized by heating *in vacuo* at about 1200° C for 24 h. Specimens afterwards taken from different parts of each respective melt (but not from the original surface) showed the same magnetic characteristics, confirming that each melt was homogeneous.

The analyses of these three alloys were:

	C1	C2	C3
carbon	0·004 7	0·001 7	0·003 3
silicon	1·78	1·70	1·40
aluminium	0·02	0·04	0·02
nickel	0·01	0·01	0·01
copper	0·01	0·01	0·01
tin	less than 0·05% for all three		
manganese	less than 0·01% for all three		
chromium	less than 0·01% for all three		
molybdenum	less than 0·01% for all three		
vanadium	less than 0·01% for all three		
cobalt	less than 0·01% for all three		
tungsten	less than 0·02% for all three		
titanium	less than 0·02% for all three		
lead	less than 0·02% for all three		

The carbon and silicon contents were measured chemically. The values for the other elements are approximate ones obtained spectrographically.

SUSCEPTIBILITY MEASUREMENTS

Susceptibilities were measured at temperatures from the Curie point upwards to above 1400° C using a ring balance[8] in a form similar to that described by Sucksmith and Pearce.[9]

The specimens were cylinders 5 mm in diameter containing an axial hole of diameter 2 mm. Generally they weighed about 250 mg. Within the apparatus they rested coaxially with and on a small projection at the end of a long vertical tube of thin-walled transparent silica (2 mm external diameter) which acted as the specimen carrier.

Temperature measurements were effected by means of a Pt/13% Pt–Rh thermocouple led inside the 2 mm tube, with its hot junction at the centre of the specimen. The thermocouple was of 36 or 42 s.w.g. wire for measurements below 1100° C, but only 36 s.w.g. was used for higher temperatures. Thermal contact with the specimen was improved by packing the space round the hot junction with a number of very small balls of pure platinum. To protect the thermocouple at the highest temperatures, the lower end of the 2 mm tube was sealed. No effects arising from contamination of the thermocouple by decomposition of the silica were observed, probably because reducing conditions were not present. As a precaution, the hot junction of the thermocouple was renewed after every two or three runs at the highest temperatures.

The ring balance is not absolute, and all susceptibilities were measured with reference to the susceptibility of pure γ-iron in the temperature range 1000 to 1050° C, taking the measurements of Sucksmith and Pearce[9] as standard. Since the phase analysis requires only comparative susceptibilities, this was quite adequate, in spite of the uncertainty caused by the small difference between the results of Sucksmith and Pearce[9] and Fallot.[10] The standard specimen used was of B.I.S.R.A. high-purity iron AH.

RESULTS

Graphs of the reciprocal of the susceptibility plotted against temperature for two iron–silicon alloys containing 2·12% and 1·04% of silicon are shown in Fig. 2.

The line *a* in Fig. 2 is characteristic over most of its length of a single-phase alloy in the α-region. The slight deviation near 1200° C results from the production of a small amount of the γ-phase, at the very tip of the γ-loop.

In the alloy represented by graph *b* in Fig. 2, the α-phase exists only below about 950° C and above about 1320° C (where it is also called the δ-phase). Between these two temperatures the susceptibility curve is mostly characteristic of the γ-phase. In this graph some hysteresis is apparent in the α- to γ-phase change, and there was an effect due to specimen evaporation near 1300° C for which a correction had to be applied.

Increasing silicon contents intermediate between these two gave a gradual transition from one type of graph to the other. First the width of the single-phase γ-region shrank and the temperature range of the mixed phase region grew, until the straight line characterizing the γ-phase failed to appear. Further increase of silicon content then gave rise merely to a deviation from the α-phase straight line which diminished as the tip of the γ-loop was approached.

The measurements led to a knowledge of the amounts of each phase present under two-phase conditions. It was a simple matter to extrapolate the straight-line dependence of the reciprocal of the susceptibility on temperature for each phase to temperatures within the mixed phase regions, giving the susceptibilities of the individual phases in any mixture. Observation of the susceptibility of the alloy at any temperature within a two-phase region then led directly by the method indicated in an earlier paragraph to an estimate of the proportions present for each phase. It was found by comparing the susceptibilities of single-phase alloys at

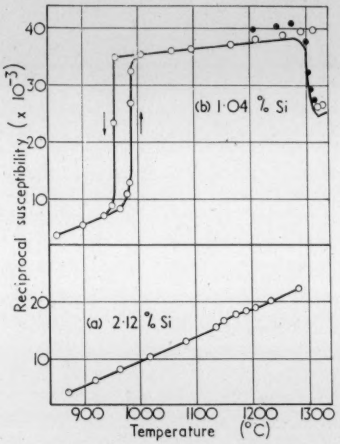


Fig. 2. Graphs of reciprocal susceptibility against temperature for two of the alloys. In graph *b* the black points were obtained during cooling from the highest temperatures. The full line in this region has been corrected for specimen evaporation

chosen temperatures with those for pure iron at the same temperature, that differences in silicon content within the range considered did not cause significant changes in the susceptibility of either phase. In the alloys for which the single-phase γ-state is never reached, it was assumed that the γ-phase had the same susceptibility as that in the 1·04% alloy. The results obtained from the analysis do not depend critically on this assumption, because, in general, the susceptibility of the γ-phase is very much lower than that of the α-phase at the same temperature.

Because temperature hysteresis in the phase change was observed, the method for recognizing equilibrium conditions was used, and Fig. 3 shows the graphs obtained for each of the alloys to which the technique was applicable. It is not clear why the locus of the maxima of the curves obtained during slow cooling does not exactly coincide with that for the minima in those for slow heating; the equilibrium lines drawn are those which give the best fit between both.

Extrapolation of the equilibrium lines to 0% and 100% γ-phase gives the temperatures of the phase boundaries for each composition. Thus a complete line for the outer boundary of the γ-loop (corresponding to 0% γ-phase) can be drawn on the phase equilibrium diagram, and also for the inner line in the range of alloys which transform completely to the single-phase γ-state. The position of the inner line can be obtained for a wider range by application of the well-known rule for the proportions of phases in a two-phase binary alloy, since the course of the outer boundary and the total composition and proportions of the two phases as a function of temperature for each alloy are now known.

The use of the method at the temperatures of the upper

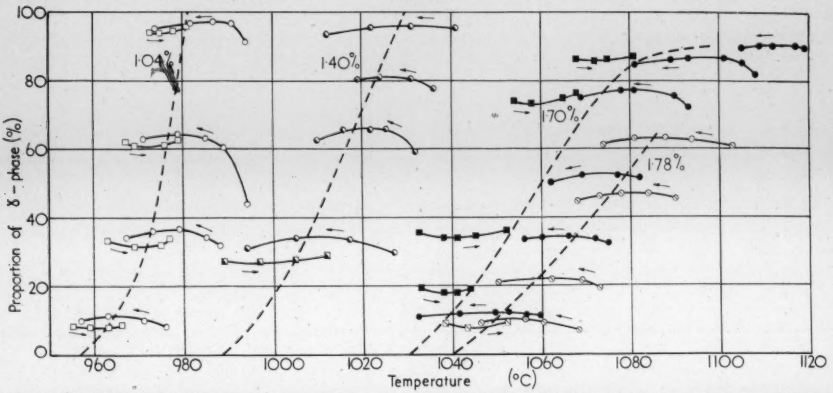


Fig. 3. Variation with temperature of the proportion of γ-phase in various iron-silicon alloys during the determination of equilibrium conditions. The broken lines represent the temperature variation of the equilibrium proportions for each alloy

half of the γ-loop was prohibited because effects attributable to the loss of active silicon from the specimens were observed during the long period of heating necessary; and also in some cases there was specimen evaporation. At these temperatures only rough estimates of the phase boundaries were made from the temperature at which the observed graphs of reciprocal susceptibility against temperature deviated from their single-phase form. Errors introduced in this way are to some extent compensated by a more rapid approach to equilibrium than at lower temperatures, but even so it is estimated that here the reliability of the points obtained is not better than about ± 20° C.

In the case of pure iron, the temperature measured for the γ- to δ-phase change was between 1380 and 1382° C, agreeing well with the previous magnetic determinations of 1385° C by Fallot[10] and between 1385 and 1395° C by Sucksmith and Pearce.[9] Here, of course, errors due to composition changes do not apply. The lines on the equilibrium diagram obtained by these methods are given in Fig. 4.

It may be significant that the results of Esser and Oberhoffer,[4] Ruer and Klesper[5] and Wever and Giani,[6] obtained using measurements on bulk properties of the materials, all tend to agree fairly well with the present work, while those of Oberhoffer and Kreutzer,[2] obtained from the use of a surface property, do not. It is to be expected that any error arising from the loss of active silicon by slight oxidation during prolonged heating in even a high vacuum will be the most marked in surface measurements. It would have the effect of displacing the apparent phase boundaries towards higher silicon contents.

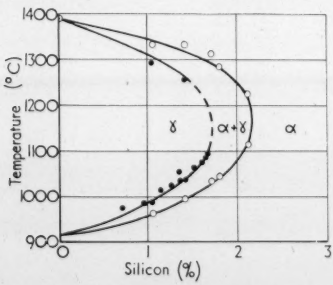


Fig. 4. The γ-loop obtained in the present work. The broken portion of the inner line is obtained by extrapolation only

ACKNOWLEDGEMENTS

The author wishes to record his gratitude to Professor W. Sucksmith for all his help. Thanks are also due to the British Iron and Steel Research Association for financial support and for supplying some pure iron; and to Dr. N. P. Allen of the National Physical Laboratory for the gift of some pure iron-silicon alloys. The chemical analyses, except for carbon, were carried out by arrangement with the British Iron and Steel Research Association at the Naval Ordnance Inspection Department, Bragg Laboratory, Sheffield. The carbon determinations were kindly undertaken by Mr. J. E. Wells at the Steel Company of Wales Ltd., Newport.

REFERENCES

(1) HANSON, M. *Der Aufbau der Zweistofflegierungen* (Berlin: Springer, 1936).
(2) OBERHOFFER, P., and KREUTZER, C. *Arch. f.d. Eisenhüttenwesen*, **2**, p. 449 (1929).
(3) MARTINDALE, R. G., and LANGFORD, D. A. *Proc. Instn Elect. Engrs*, **100** (Pt II) p. 417 (1953).
(4) ESSER, H., and OBERHOFFER, P.* *Ber. d. Fachaussch. d. V. d.Eisenhütt., Werkstoffausschuss*, No. 69 (1925).
(5) RUER, R., and KLESPER, R.* *Ferrum*, **11**, p. 257 (1913).
(6) WEVER, R., and GIANI, P.* *Mitt. K.-Wilh.-Inst. Eisenforsch.*, **7**, p. 59 (1925).
(7) GREINER, E. S., MARSH, I. S., and STOUGHTON, B. *The Alloys of Iron and Silicon* (New York: McGraw-Hill Book Co. Inc., 1933).
(8) SUCKSMITH, W. *Phil. Mag.*, **8**, p. 158 (1929).
(9) SUCKSMITH, W., and PEARCE, R. R. *Proc. Roy. Soc. A*, **167**, p. 189 (1938).
(10) FALLOT, M. *J. Phys. Radium*, **5**, p. 153 (1944).

* See Greiner, Marsh and Stoughton[7].

New books

Measurement techniques in mechanical engineering. By R. J. SWEENEY. (London: Chapman and Hall, Ltd., New York: John Wiley and Sons, Inc.) Pp. x + 309. Price 44*s*.

This book has been written with the intention of providing a source of information on the measurement techniques which are commonly used in performance testing of power equipment. The scope of the book is therefore rather more restricted than the use of the phrase "mechanical engineering" in the title would suggest; the fields of materials testing and experimental stress analysis are specifically excluded by the author from consideration in this book.

The book deals with the measurement of length, area, time, speed, force, pressure, temperature and fluid flow rates. There are also brief discussions of elementary electrical methods of measurement and a few examples of chemical methods in common use in this field. A book of this sort might easily be just a catalogue of instruments in common use. This book is much more than that. The author discusses the principles involved in the design of the instruments and each of the major chapters contains something in the nature of a critical review of the measuring techniques under discussion. The chapters on temperature measurements and rates of fluid flow are particularly well balanced. The book clearly does not set out to be an encyclopaedia, but it seems surprising that more attention has not been paid to the use of variable capacitance and variable inductance units for pressure measurements, or the use of resistance strain gauges in pressure cells; also only one method of gas analysis is described.

The author has, in fact, succeeded in doing what he set out to do and has provided, within the confines of one attractively produced book, a great deal of the information about measuring techniques which engineers concerned with the testing of power equipment may have to use.

E. K. FRANKL

Radio engineering, Vol. 1. 2nd edition. By E. K. SANDEMAN, Ph.D., A.C.G.I., M.I.E.E. (London: Chapman and Hall Ltd.) Pp. xxiv + 779. Price 60*s*.

The first editions of Vols. 1 and 2 were reviewed in the *Journal of Scientific Instruments* in April 1950 and the 2nd edition of Vol. 2 in the November 1953 issue of the *British Journal of Applied Physics*. In the 2nd edition of Vol. 1, the opportunity has been taken to make a number of corrections and to revise the data concerning the details of broadcasting practice. The text as a whole, and the numbering of the sections, remain unchanged. T. B. RYMER

Introduction to a study of mechanical vibration. By G. W. VAN SANTEN. (London: Cleaver Hume Press Ltd.; Eindhoven: Philips Technical Library.) Pp. xvi + 296. Price 35*s*.

To survey the vast subject of mechanical vibrations, in a volume containing 288 pages of text, at the beginning of which not even the properties of a sine wave are assumed, is a remarkable achievement. Writing under such circumstances simplification is inevitable, but the author's presentation of the basic ideas is good and aided by many useful diagrams; Chapter XV, dealing in elementary terms with aircraft flutter, vibration of cables, wheel wobble, etc., is particularly successful. The book is not intended to be mathematical but, when appendices are included which list the simplest trigonometric formulae and define complex numbers, a short bibliography of helpful elementary mathematic texts is surely desirable, especially as other references are plentiful. The first chapter would be greatly improved by rearrangement so that the ubiquity of the harmonic equation in vibration theory and the consequent interest in its solutions sine and cosine are mentioned before their properties are discussed; otherwise the sequence is good.

Later chapters are devoted to description of instruments for vibration measurements; the principles used are clearly stated and many photographs of commercial equipment are given: of these the products of the author's own firm, who are also the publishers, naturally tend to predominate.

A short list of errata is included but unfortunately does not yet cover all the typographical errors, some of which, such as the omission of π from equation (105) on p. 39, could cause slight confusion. The volume is nicely produced and the English translation reads naturally. M. J. P. MUSGRAVE

Statistical astronomy. By R. J. TRUMPLER and H. F. WEAVER. (London: Cambridge University Press; Berkeley and Los Angeles: The University of California Press, 1953.) Pp. xx + 644. Price 56*s*. 6*d*.

The first two hundred pages of this book give a clear account of the elements of modern statistical theory: such subjects are discussed as univariate and multivariate frequency functions, integral equations of statistics, general theory of samples, statistical uncertainty; methods of testing hypotheses, Student's t-test and the χ^2-test are also examined; examples are worked out, such as an application to the motions of a particular class of stars of the χ^2-test; in the case discussed it leads to the rejection of the ellipsiodal theory of star-streaming. Later in the volume the ellipsiodal theory is shown to give the best statistical model for the motions of stars in the vicinity of the sun.

In a general statistical examination of the galaxy the luminosity-spectral type distribution is analysed and a series of special studies of the Hertzsprung–Russell diagram are given. The space distribution of stars is discussed and reasons given for the abandonment of the Kapteyn Universe. Galactic rotation and the two-population of stars worked out by Baade are among the factors contributing to the solution of many problems in statistical studies of galactic structure. The long experience of Professor Trumper in this field of work, coupled with the recent researches of Professor Weaver, has produced a mine of information and a thorough and well-balanced study of galactic structure that no one working in the subject can do without. Each section of the book is followed by a most useful bibliography.

F. J. M. STRATTON

Dimensional methods and their applications. By C. M. FOCKEN, B.Sc., D.Phil., M.S., F.Inst.P. (London: Edward Arnold and Co.) Pp. viii + 224. Price 30*s*.

It is curious that the method of dimensions has led to the publication of so many papers and, in consequence, to so much disagreement. Anyone interested in the why and wherefore of the method and the ways in which it can be used (for it does work in spite of the disagreement) might therefore well be pleased to read an account broadly based on the literature. Dr. Focken's book is such an account, and a good account, too. Its defects are defects of the literature and may be summarized and exemplified by the phrase which occurs several times in it: "it is the author's opinion." E. W. H. SELWYN

Notes and comments

Elections to The Institute of Physics

The following elections have been made by the Board of The Institute of Physics:

Fellows: W. H. Denton, P. J. Duncton, A. J. Ede, W. J. Lewis, G. D. Morgan, H. Potter, C. A. Taylor, R. S. Tebble, C. Wood.

Associates: W. H. Atkinson, P. J. Becque, B. Birtwistle, R. A. Bones, D. W. G. Byatt, C. A. Clark, R. H. Creamer, R. W. Dudding, H. W. Emerton, W. A. Evans, P. L. Flowerday, D. F. Halliday, P. Halliday, C. M. Hargreaves, C. W. Harland, R. Haynes, F. J. Hiorns, J. R. Hodkinson, J. Ingham, R. T. Jarman, D. W. Jones, P. Lord, R. W. Mackay, P. F. Mariner, D. H. Martin, D. R. Milner, H. B. Mohanti, D. Murray, G. B. F. Niblett, P. R. D. Pomeroy, D. Raynor, W. T. Roberts, J. Sikorski, J. J. Sparkes, R. W. Taylor, J. A. Thomas, D. T. Turnbull.

Fifty-five Graduates, twenty-eight Students and five Subscribers were also elected.

Recommendations of the International Commission on Radiological Units

The International Commission on Radiological Units was originally set up to deal with the physical aspects of dosage in medical radiology. Developments of recent years, when a larger proportion of the population has become radiation conscious, have greatly widened the field of application of its recommendations. The first part of the recent revision of these recommendations, carried out in Copenhagen last year and dealing with definitions and units, will be of interest to physicists beyond the field of medicine.

Although its range of practical usefulness is limited to quantum energies up to 3 MeV, the roentgen is retained as the unit of X- or γ-ray dose, with the definition unchanged. By far the most interesting development is the adoption of the term *absorbed dose* of any ionizing radiation as the amount of energy imparted to matter by ionizing particles per unit mass of irradiated material at the place of interest. The *rad* is named as the unit of absorbed dose and is 100 ergs per gramme. The adoption of this unit should help to overcome the confusion which is liable to arise in radiation physics and radiobiology when dealing with the wider range of ionizing radiations. To facilitate the conversion of ionization measurements to energy absorption, the Commission proposes to publish tables of data on the relevant parameters and also of radioactive constants.

The revised recommendations are published in the *British Journal of Radiology*, Vol. XXVII, April 1954.

Second annual meeting of ASTM Committee on Mass Spectrometry

The second annual meeting of ASTM Committee E-14 on Mass Spectrometry will be held from 24 to 28 May, 1954, at the Jung Hotel, New Orleans, Louisiana. Forty-six papers will be presented covering isotopic abundance determinations, ionization potentials, solids analysis, high molecular weight analysis, and new mass spectrometer techniques and developments.

All those interested are invited to attend. Further details may be obtained from Mr. William Priestley, Jr., Standard Oil Development Company, P.O. Box 121, Linden, New Jersey, U.S.A.

International convention on oxide coated cathodes

The Société Française des Ingénieurs Techniciens du Vide is organizing an international convention on oxide coated cathodes to be held in Paris on 24–25 June, 1954. Papers should be submitted before 15 May, 1954

Further information may be obtained from The Société Française des Ingénieurs Techniciens du Vide, 44 rue de Rennes, Paris VI°.

High vacuum symposium

It is announced that a high vacuum symposium will be held at the Berkeley Carteret Hotel, Asbury Park, New Jersey, U.S.A., on 16–18 June, 1954. Papers for presentation will be divided into five major groups: (1) nomenclature and standards; (2) new equipment and instruments; (3) fundamental developments in vacuum technology; (4) methods and techniques; (5) applications and processes.

The Honorary Chairman of the symposium will be Dr. A. C. Hickman and further details may be obtained from the Committee on Vacuum Techniques, P.O. Box 1282, Boston 9, Massachusetts, U.S.A.

Journal of Scientific Instruments

Contents of the April issue

ORIGINAL CONTRIBUTIONS

Papers

The behaviour of standard cells in conditions which include the generation of appreciable current. By L. Hartshorn and Freda Manning.

Effects of ambient-temperature variations on glow-discharge tube characteristics. By F. A. Benson and G. Mayo.

Spatial variations of the spectral response of photomultiplier cathodes. By H. Edels and W. A. Gambling.

A visual tricolorimeter using the C.I.E. stimuli X, Y, and Z. By R. W. G. Hunt.

An r.m.s. milliammeter of novel design for the measurement of current from zero to video frequencies. By H. B. Wood.

A portable photometer for the testing of reflective road signs. By R. G. Giovanelli.

A thermistor McLeod gauge for a pressure range, $1–10^{-7}$ mm of mercury. By R. S. Bradley.

A precision 19 cm X-ray diffraction camera. By J. Adam.

Inhomogeneity e.m.f.s in thermoelectric thermometers. By N. Fuschillo.

A sealed-off metal proportional counter for cosmic-ray work. By J. S. Buchanan.

A fully automatic recording densitometer for scanning paper electrophoresis patterns. By D. J. R. Laurence.

A servo-system for accurate speed control. By L. U. Hibbard, D. E. Caro and J. Y. Freeman.

LABORATORY AND WORKSHOP NOTES

A simple mercury cup connexion for Geiger counters. By A. H. Ward.

A magnetic mercury cut-off. By F. W. Thompson.

An apparatus for delivering vapour of constant composition and pressure. By A. Charnley, G. L. Isles and J. S. Rowlinson.

Relay unit for time signal receiver. By J. Hers.

A tap for corrosive vapours. By A. B. Osborn.

A simple tool for handling small specimens. By S. P. Anderson.

Some techniques for making stable non-rectifying contacts to germanium and other semi-conductors. By W. H. Mitchell.

NOTES AND NEWS

Correspondence

A mould for pressing plastic disks. From R. Hayes.

Mechanical precision required in electrostatic analysers. From M. Hoyaux and J. Geets.

New books

New instruments, materials and tools

Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2*s.* 6*d.* including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

The acceleration of heavy ions in a fixed-frequency cyclotron*

By D. WALKER, M.Sc., Ph.D., F.Inst.P., J. H. FREMLIN, M.A., Ph.D., F.Inst.P., W. T. LINK, M.Sc., and K. G. STEPHENS, B.Sc., Physics Department, University of Birmingham

A study has been made of the acceleration of the ions $^{9}Be^{4+}$, $^{12}C^{5+}$, $^{13}C^{5+}$, $^{13}C^{6+}$, $^{14}N^{6+}$, $^{15}N^{6+}$, $^{16}O^{6+}$ and $^{20}Ne^{9+}$ in a 60 in. fixed-frequency cyclotron. Since the momentum which can be imparted to an ion in a given cyclotron magnet is proportional to the charge carried by the ion, one of the chief problems in accelerating heavy ions to high energies is that of obtaining highly-charged ions. Of particular interest are ions with a final charge which is a multiple of 3, as ions with one-third of the final charge can be extracted from the ion-source and accelerated at closely the same value of magnetic flux-density as the final ion. These ions with one-third of the final charge have a frequency of revolution which is one-third that of the final ions. This preliminary "third harmonic" acceleration of the ions with a low charge gives them such an energy that collisions with gas atoms in the cyclotron tank can strip the ions to the final state of charge desired. This mechanism of acceleration leads to a continuous energy-distribution of the final ions of high charge. The beam current of high energy ions can be increased by increasing the pressure of gas in the cyclotron tank to increase the probability of stripping.

1. INTRODUCTION

The high-energy ions which have claimed the greatest attention from nuclear physicists are those of the hydrogen and helium isotopes. However, fast ions of heavier elements are also of considerable interest. Such ions were first observed in nuclear disintegration and recoil processes, and in the fission of heavy nuclei. Later, in 1948, atomic nuclei of moderate mass were observed to be present in the primary cosmic radiation.[1] The acceleration of a heavy ion ($^{12}C^{6+}$) in a cyclotron was first reported in 1940[2] but, although since then several investigations using heavy ion beams of high energy have been reported, the subject has received relatively little attention until recently. Several investigations of nuclear reactions produced by heavy ions have been recorded.[3-9] Radio-chemical, counter, and photographic emulsion techniques have been employed in studying these reactions, but there is undoubtedly still much information to be gained. Such reactions may prove useful for particular purposes, such as the production of transuranic elements where the product nucleus can differ greatly in mass and charge from an available target nucleus. Breit, Hull and Gluckstern[10] have suggested that the bombardment of nuclei with heavy ions can yield new information on nuclear structure. Other investigations with cyclotron-accelerated heavy ions have been concerned with the variation in the charge carried by the ions as they are slowed down in their passage through matter,[11] with the density of δ-rays (recoil electrons) along the tracks of heavy ions in photographic emulsions[6] and with the multiple Coulomb scattering of the ions in emulsions.[6]

The first successful operation of a cyclotron was reported by Lawrence and Livingston as long ago as 1931. The slowness with which the cyclotron has since been applied to the acceleration of ions heavier than ^{4}He can be attributed to two causes. Not only was the bombardment of atomic nuclei with the simplest possible particles best calculated to lead to a rapid advance in our understanding of the properties of nuclei, but the actual acceleration process involves, for heavy ions, certain technical complications which are absent in the case of very light ions. It is with the mechanism of acceleration of heavy ions that we are concerned in the present paper, which describes some investigations carried out with the 60 in., fixed-frequency cyclotron in the Nuffield Laboratory at the University of Birmingham. Although these technical studies are not exhaustive, they have elucidated some important effects.

* The data in this article were presented on 18 July, 1953, by D. Walker at the Conference on Nuclear Physics held at the University of Birmingham.

2. GENERAL FACTORS GOVERNING THE ACCELERATION OF HEAVY IONS IN A CYCLOTRON

Before proceeding to a discussion of investigations carried out with the Birmingham cyclotron, we shall note some general points concerning the acceleration of heavy ions in a cyclotron.

The cyclotron method of acceleration is applicable to ions of any mass-to-charge ratio. The frequency of revolution (f) of an ion in a plane at right angles to a magnetic flux is given by

$$2\pi f = B(q/M_0)(1 + E/M_0c^2)^{-1} \qquad (1)$$

where B is the magnetic flux-density, q the charge carried by the ion, M_0 its rest mass and E its kinetic energy. To satisfy the conditions for cyclotron acceleration it is only necessary that the frequency of the alternating accelerating voltage on the "dees" of the cyclotron be set equal to f or to an odd multiple of f. Clearly an appropriate combination of B and radio-frequency can always be found for any ion.

Apart from the effect at high energies of E in equation (1), B/f is proportional to M_0/q. If we adopt the electronic charge as our unit of charge and the atomic mass unit as our unit of mass, then M_0/q is very nearly equal to 1 for the proton, 2 for the deuteron and 3 for the triton. For the hydrogen isotopes these are the only values possible, but for isotopes of other elements there exist as many values of M_0/q as there are possible states of ionization. In the case of ^{3}He for example, the possible values are 3 and 1·5, while in the case of ^{16}O, the possible values are 16, 8, 5·33, 4, 3·2, 2·66, 2·29 and 2. It is worth noting that, for all stable isotopes heavier than ^{3}He, M_0/q is greater than in the case of the deuteron (except for fully-stripped nuclei containing equal numbers of protons and neutrons). Thus, with the exceptions noted, B/f for heavy ions is always greater than in the case of the deuteron.

The charge carried by a heavy ion has another important effect. The energy which can be imparted to an ion of a

particular mass, in a given cyclotron magnet, increases with the ionic charge. This follows from the relation

$$p = Bqr \qquad (2)$$

where p is the momentum of an ion of charge q at a radius r where the magnetic flux-density is B. If it is desired to produce an ion of the highest possible energy then q must be as large as possible. On the other hand the binding energy of atomic electrons increases with atomic number and it becomes increasingly difficult, the heavier the atom, to produce high states of ionization directly in or near the ion-source of the cyclotron. This is an important technical feature in a comparison of the acceleration of light and of heavy ions. The effect is actually already apparent if we compare the cyclotron acceleration of hydrogen and helium isotopes. It is well known that the beams of doubly-charged helium which can be obtained are weaker than the corresponding beams of singly-charged deuterium (see, for example, University of Birmingham Cyclotron[12]). It is interesting that a 63 in. cyclotron has been constructed at Oak Ridge National Laboratory for the acceleration of $^{14}N^{3+}$ ions.[8] This represents a compromise giving a moderate final energy (25 MeV) and a relatively large current of ions (1 μA).

Another feature which is peculiar to the acceleration of ions heavier than helium is that not all elements can be obtained in gaseous compounds suitable for use in the usual gas-discharge ion-source of cyclotrons. Thus in some cases, as for example ^{9}Be, solid sources of ions must be used.

3. METHODS OF OBSERVING BEAMS OF HEAVY IONS INTERNALLY IN THE CYCLOTRON

In studies of the acceleration of heavy ions in the Birmingham cyclotron we have used four distinct methods for detecting internal beams of ions. In each case the sensitive element was part of a probe which could be inserted through an air-lock into the cyclotron vacuum and set at a suitable radius in the gap between the "dees." The sensitive elements which have been used are: a current-collector, a thermocouple, a scintillating crystal, and an Ilford Nuclear Research photographic plate. We shall describe each of these arrangements briefly.

(*a*) *Current-collecting probe.* The inner end of the current-collecting probe is shown in section in Fig. 1. The standard

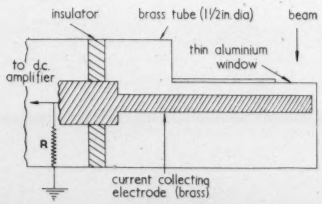


Fig. 1. Current-collecting probe

1½ in. diameter brass tube which slides into the cyclotron vacuum between the "dees" has at the end a thin aluminium window facing the direction of the oncoming ion beam. An insulated brass rod on the axis of this tube is attached to a brass plate in which the beam is arrested. A resistor R (up to $10^{11}\,\Omega$) is connected between the rod and tube outside the cyclotron vacuum and the potential drop across this resistor is amplified by means of a d.c. amplifier. We have found this probe to be convenient for measuring currents of about 10^{-11} A upward. For smaller currents it becomes difficult, without special precautions, to guard against the strong electrical interference which is inevitably present in and near the cyclotron. The device is not suitable for the measurement of resonant heavy ion beams of very low energy as it is found that the aluminium window-foil cannot be removed without recording comparable currents of low-energy, non-resonant ions. As an example, an aluminium window 5 μ thick might be used, in which case we could not record beams of nitrogen ions below about 7 MeV in energy. The probe is most suited to the measurement of higher-energy beams of suitable intensity. For the most precise measurements of beam intensities, the secondary emission of electrons from the surface struck by the beam should be taken into account. However, the effect of such secondary emission must be small in the presence of the intense magnetic field of the cyclotron, and in any event, the highest precision was not necessary in the investigations described below. For the highest precision the exact charge state of the heavy ions at the moment of collection would also have to be known.

(*b*) *Thermocouple probe.* The current-collecting probe just described has a similar sensitivity for ions of all energies. It is also useful to have a probe which is by contrast sensitive to the total beam power which it intercepts. We have used a sensitive thermocouple element in such a probe, which is shown diagrammatically in Fig. 2. The ion beam is stopped

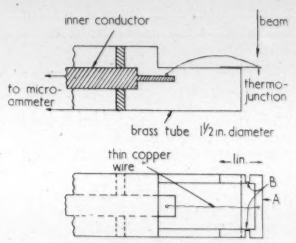


Fig. 2. Thermocouple probe

by a strip of Ferry metal (a 45% nickel–copper alloy made by Henry Wiggin and Co. Ltd.) A which is connected by thin strips of Ferry metal B to the main 1½ in. brass tube. To the middle of the strip of Ferry metal is silver-soldered a thin copper wire about 3 in. long which connects with the inner conductor on the axis of the brass tube. Externally to the cyclotron a microammeter is connected between the brass tube and the inner conductor. Such a probe has been used to record beam powers of the order of a watt.

(*c*) *Scintillating crystal probe.* For the most sensitive detection of heavy ion beams a detector of individual high-energy particles cannot be surpassed. If such a detector has an energy-dependent response, an investigation of the energy-distribution of the particles can readily be made. An energy-dependent response is, in any case, essential if the particle-detector is to be in or near a cyclotron in the presence of an inevitably high background of low-energy radiation. We have found a scintillating crystal of potassium iodide activated with thallium to be very suitable as such a detector. For use inside the cyclotron, the crystal is mounted on the end of a ¾ in. diameter perspex light guide, as shown in Fig. 3, and

inserted into the cyclotron vacuum between the "dees." The light guide leads to a photomultiplier situated outside the cyclotron in the relatively weak fringing magnetic field where it is covered with a Mumetal screen. With strong internal beams of ions the probe shown in Fig. 3 has sometimes been

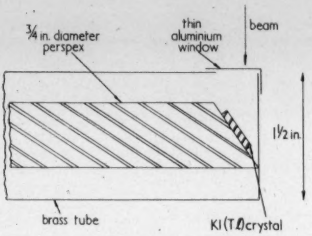


Fig. 3. Scintillating crystal probe

used in modified form designed to restrict the ions incident on the crystal to a finely collimated pencil by means of a suitable disposition of slits and crystal.

(*d*) *Photographic plate probe.* Another method of detecting beams of heavy ions, which has been used to some extent internally as well as externally to the cyclotron, has been that of allowing the beam to fall on a Nuclear Research photographic plate. For internal use the plates can be mounted on the end of a suitable probe and inserted between the "dees." Although a rather slow method of gathering data in a technical study of acceleration processes, the use of photographic plates has been useful in confirming the nature of the particles accelerated.

4. MAGNETIC RESONANCE CURVES FOR HEAVY IONS

As has been mentioned already, the process of accelerating heavy ions to high energies in a cyclotron has some complicating features when compared with the acceleration of the light hydrogen or helium ions. It is best, therefore, to precede a discussion of the magnetic resonances of highly-charged heavy ions with a consideration of the resonance curve of a heavy ion carrying a low charge. A good example to choose is the $^{14}N^{2+}$ ion.

(*a*) *Magnetic resonance curve for* $^{14}N^{2+}$ *ions.* The curve on the left in Fig. 4 was obtained by using the sensitive thermocouple probe at a radius of 20 in. The ion-source was fed with nitrogen gas. The flux-density* at which this resonance is observed is such that the frequency of revolution [f in equation (1)] of the $^{14}N^{2+}$ ions is one-third of that of the radio-frequency accelerating voltage. (We call ion acceleration under these conditions "third harmonic" acceleration.) At a radius of 20 in. the energy of the $^{14}N^{2+}$ ions is about 9 MeV. The peak thermocouple current recorded in this particular measurement was estimated to correspond to an incident ion current of about 0·2 μA. This ion current was very much larger than that of the higher energy $^{14}N^{6+}$ ions which were present simultaneously and consequently the form of the resonance curve is wholly determined by the $^{14}N^{2+}$ ions.

(*b*) *Magnetic resonance curves for highly-charged heavy ions.* The curve on the right in Fig. 4 is a magnetic resonance curve for $^{14}N^{6+}$ ions obtained by using the scintillating crystal probe at a radius of 24 in. At the $^{14}N^{6+}$ resonance, $^{14}N^{2+}$ ions which have undergone third harmonic acceleration will also be present. The $^{14}N^{2+}$ ions do not, however, contribute to the observed counting rates, since the aluminium window of the scintillating crystal probe was deliberately made thick enough to stop them completely.

The displacement of the $^{14}N^{6+}$ (24 in. radius) and $^{14}N^{2+}$ (20 in. radius) resonance curves in Fig. 4 is partly due to the manner in which the flux-density is plotted [see footnote to Section 4(*a*)]. It is estimated that, if the resonance curves

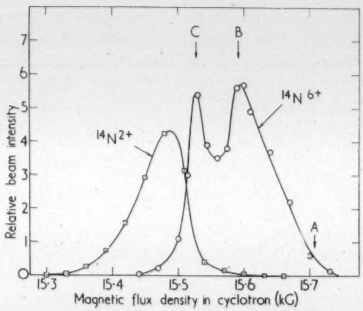


Fig. 4. Magnetic resonance curves for $^{14}N^{2+}$ and $^{14}N^{6+}$ ions at 20 and 24 in. radius respectively. The frequency of alternating accelerating voltage in the cyclotron is 10·28 Mc/s. The scale of relative beam intensity relates independently to the two resonances. The maximum intensity of the $^{14}N^{2+}$ beam is actually very much greater than that of the $^{14}N^{6+}$ beam

had been measured at the same radius, the peak of the $^{14}N^{2+}$ resonance curve would have coincided approximately with that peak C of the $^{14}N^{6+}$ resonance which occurs at the lower value of flux-density. It should be noted that, at half-maximum, the $^{14}N^{6+}$ curve is much broader than the $^{14}N^{2+}$ curve.

Resonance curves similar to that of $^{14}N^{6+}$ have been observed also with $^{13}C^{6+}$, $^{15}N^{6+}$ and $^{16}O^{6+}$ ions. While the occurrence of two peaks in such resonances appears to be very characteristic, the exact form of the resonance and the relative magnitudes of the two peaks is very dependent on such factors as the amplitude of the radio-frequency accelerating voltage on the "dees," the precise positions of the

* A word of explanation is required about the flux-density scales of Figs. 4–6. The flux-density plotted is, for the range of flux-density in any one figure, a linear function of the current flowing in the magnet coils. The relation of the plotted flux-density to the actual flux-density in the cyclotron is such that, if the magnetic resonance curve for deuterons were plotted by using a detector at a radius of 24 in., the plotted flux-density at the peak of the resonance (B_p) would satisfy the relation $2\pi f_r = B_p(e/M_d)$ where f_r is the frequency of the alternating accelerating voltage and e/M_d the charge-to-mass ratio of the deuteron. Since the actual flux-density in the cyclotron falls slowly with distance from the centre of the cyclotron (in order to provide vertical focusing of the beam), the average flux-density encountered by an ion accelerated from the centre of the cyclotron to a detector depends slightly on the radius at which the detector is situated.

"dees" and the radius of detection. We shall not therefore discuss here the structure exhibited by these resonances or its possible significance. The occurrence of structure is a complex phenomenon which may be peculiar to a particular cyclotron. It does not bear on the conclusions of this paper. To illustrate the dependence of the shape of $^{14}N^{6+}$ resonance on the operating conditions, we show in Fig. 5 two resonance curves taken under identical conditions except that in the case of the upper curve the right "dee" has been moved about 1 cm nearer the ion-source. The lower curve is the same as that in Fig. 4.

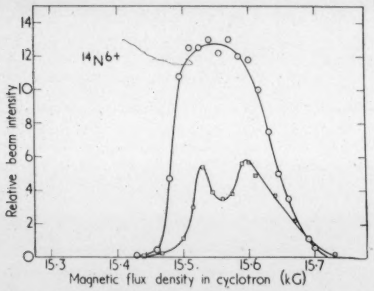


Fig. 5. Magnetic resonance curves for $^{14}N^{6+}$ ions (24 in. radius) for two slightly different positions of the right-hand "dee" of the cyclotron

Fig. 6 is a magnetic resonance curve for $^{12}C^{5+}$ ions obtained by using the scintillating crystal probe at a radius of 24 in. The ion source was fed with carbon dioxide. At the $^{12}C^{5+}$ resonance it is possible that $^{12}C^{1+}$ ions which have undergone fifth harmonic acceleration will be present. Again these $^{12}C^{1+}$ ions do not contribute to the observed counting rates of Fig. 6, since they are completely stopped by the aluminium

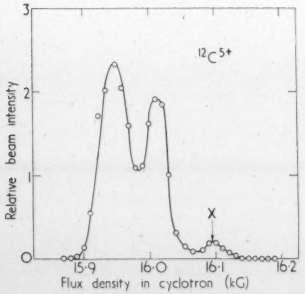


Fig. 6. Magnetic resonance curve for $^{12}C^{5+}$ ions (24 in. radius). The cyclotron frequency is 10·31 Mc/s

window. As in the case of $^{14}N^{6+}$, the main resonance in this example exhibits two peaks which we shall not discuss further here. Attention is drawn, however, to the small peak *X*, occurring on the high flux-density side of the resonance, to which reference will be made later at the end of Section 8.

5. LOCATION OF THE RESONANCES DUE TO VARIOUS HIGHLY-CHARGED HEAVY IONS

Some of the material in this section has already been reported,[7] but is included here for completeness.

The scintillating crystal probe has been used to locate the cyclotron resonances of heavy ions over a wide range of magnetic flux-densities. The results discussed here were obtained with the probe set at a radius of from 24 to 27 in. The window of the probe consisted of aluminium foil 17 μ thick.

Resonances due to the following ions, listed in order of increasing mass-to-charge ratio, were located: $^{13}C^{6+}$, $^{20}Ne^{9+}$, $^{9}Be^{4+}$, $^{14}N^{6+}$, $^{12}C^{5+}$, $^{15}N^{6+}$, $^{13}C^{5+}$ and $^{16}O^{6+}$. At a radius of 27 in. in the Birmingham cyclotron, the kinetic energy of an ion revolving about the centre of the cyclotron, with a frequency equal to that of the radio-frequency accelerating voltage, is approximately $10A$ MeV, where A is the atomic mass number of the ion. This maximum attainable energy thus varies from 90 MeV for $^{9}Be^{4+}$ to 200 MeV for $^{20}Ne^{9+}$. To produce carbon or oxygen ions, the ion source was fed with carbon dioxide gas; to produce nitrogen ions, with nitrogen gas; to produce neon ions, with neon gas. In order to produce beryllium ions, the usual molybdenum tip of the ion-source was replaced by one of beryllium metal and the ion source was fed with hydrogen gas. When the cyclotron was running, this beryllium tip became heated to a bright red.

In Fig. 7 the flux-density in the cyclotron is plotted against the current flowing in the coils of the magnet. Indicated on

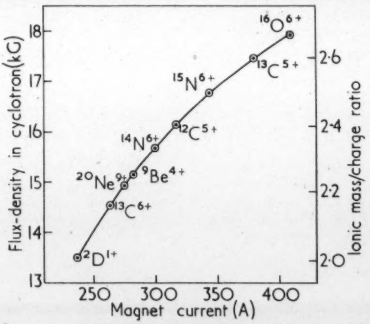


Fig. 7. Location of the resonances of highly-charged heavy ions. The cyclotron frequency is 10·31 Mc/s. Mass-to-charge ratios are expressed as (atomic mass)/(number of electron charges)

this curve, at the measured magnet currents, are the resonances due to the different ions. (The $^{2}D^{1+}$ resonance, located by means of an insensitive thermocouple probe, is included for comparison.) These resonances correspond exactly, for the cyclotron frequency of 10·31 Mc/s, to the mass-to-charge ratios given by the right-hand scale. This provides unequivocal identification of the resonances. The scintillating crystal probe has not been used to study resonances of ions

for which the mass-to-charge ratios are close to those of $^2D^{1+}$ and $^4He^{2+}$ because of interference by intense beams of the latter ions arising from contamination of the cyclotron. It has in any event been considered of some advantage to work with heavy ions, the resonances of which are widely separated from those of $^2D^{1+}$ and $^4He^{2+}$. This eliminates the possibility of experiments, which use the heavy ion beams, becoming complicated by the possible presence of $^2D^{1+}$ and $^4He^{2+}$ ions.

Harmonic pre-acceleration. During the location of the resonances due to the various ions, a striking feature noticed was the great prominence of the $^{13}C^{6+}$, $^{14}N^{6+}$, $^{15}N^{6+}$ and $^{16}O^{6+}$ resonances. The $^{13}C^{5+}$ resonance is about 700 times weaker than the $^{13}C^{6+}$ resonance while the $^{15}N^{7+}$ and $^{16}O^{7+}$ resonances are not detectable at all. A careful search for the $^{16}O^{7+}$ resonance showed it to be at least 10000 times weaker than the $^{16}O^{6+}$ resonance.

The frequency of revolution of an ion in the cyclotron is governed by equation (1). When the cyclotron is adjusted to accelerate six-charged ions, doubly-charged ions of the same atomic species will have a frequency of revolution which is one-third of the cyclotron radio-frequency. This is strictly true only in the approximation $E \ll M_0c^2$, since for a given orbit radius, doubly-charged ions will have a kinetic energy which is only one-ninth of that of a six-charged ion. In our case $E \lesssim (1/100)M_0c^2$, so that the condition is closely fulfilled and doubly-charged ions can be accelerated. There can be no doubt that pre-acceleration of ions in the doubly-charged state (third harmonic pre-acceleration) aids in an important manner the ultimate stripping of the ions to the six-charged state through collisions with gas atoms in the cyclotron tank. (The third harmonic beam is not itself detected by the scintillating crystal since this beam is stopped completely by the 17 μ of aluminium covering the crystal.) Fifth and seventh harmonic pre-acceleration are increasingly less effective for this purpose. With neon ions, only the $^{20}Ne^{9+}$ resonance, which fulfills the third harmonic condition, has so far been identified.

Condit[13] and Miller[5] have mentioned that harmonic acceleration can take place with carbon ions, but the importance of harmonic pre-acceleration for the subsequent stripping of electrons appears to have been first clearly demonstrated by the present results.[7] (Recent reports privately communicated from the Radiation Laboratory, University of California, Berkeley, show that this stripping mechanism had earlier been postulated by Hollander and others independently to explain their observations on carbon ions.)

In the case of $^9Be^{4+}$ ions, odd harmonic pre-acceleration is not possible. It is felt, however, in this case that, owing to the comparatively low atomic number of beryllium, it is relatively easy to get atoms fully stripped without recourse to harmonic pre-acceleration and so obtain an appreciable beam. Beams of $^9Be^{4+}$ have been obtained which are about as strong as the beams of $^{13}C^{6+}$ obtained when carbon dioxide of normal isotopic constitution is used in the ion-source.

6. ENERGY-DISTRIBUTION OF THE IONS

The scintillating crystal probe lends itself to a study of the energy-distribution of the ions incident on the crystal through an analysis of the pulse-height distribution observed at the output of the photomultiplier. For heavy ions incident on a KI(Tl) crystal, one would expect the relation between light output and the incident particle energy to be somewhat non-linear on account of the high ionization density along the paths of the heavy ions stopping in the crystal,[14] but for qualitative purposes the precise form of the relation is not important. The most important thing to know is whether the ions have a unique energy at a given radius or whether the ions are continuously distributed in energy. Further, it is useful to have a ready means available for studying any changes in the energy-distribution of the ions which may be brought about by changing the operating conditions of the cyclotron.

(*a*) *Energy-distribution of $^3He^{2+}$ ions.* In order to test the scintillating crystal probe on a particle spectrum which should show a monoenergetic component, a pulse-height analysis was made with a $^3He^{2+}$ ion beam, at a radius of 17·7 in., using a differential discriminator. (A $^3He^{2+}$ beam was used instead of either a $^4He^{2+}$ or a $^2D^{1+}$ beam because the cyclotron is normally used to accelerate the latter ions and contamination of the ion-source makes it difficult to obtain internal beams of these ions weak enough for counting, even when neither deuterium nor helium gas is being fed to the ion-source.) To obtain a conveniently low counting rate, the gas fed to the ion-source was a mixture of nitrogen with 2% of natural well helium, which itself contains only one part in 10^7 of ^{3}He, and the beam impinging on the crystal was restricted by fine slits. Fig. 8 shows the result. The

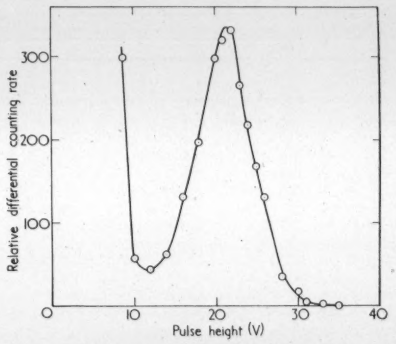


Fig. 8. Pulse-height analysis on an internal $^3He^{2+}$ ion beam with the scintillating crystal probe

energy of the $^3He^{2+}$ ions giving rise to the peak in the distribution is 13 MeV and such a monoenergetic component of relatively low energy is clearly resolved.

(*b*) *Energy-distribution of heavy ions.* Investigations of energy distributions by means of pulse-height analysis have been carried out with $^{13}C^{6+}$, $^{14}N^{6+}$, $^{15}N^{6+}$ and $^{12}C^{5+}$ ion beams at a radius of 24 in. in the cyclotron. The nature of the energy spectrum is very similar in all these cases, and depends at what point on the magnetic resonance curve [Section 4(*b*)] the cyclotron is set. The spectra are continuous, but as the flux-density is increased from low to high values through the resonance, the number of high-energy ions relative to low-energy ions increases. This is illustrated in Fig. 9 which shows some differential pulse-height analyses with a $^{14}N^{6+}$ beam. The legends *A*, *B*, *C* refer to the points indicated on the resonance curve of Fig. 4. At the point *A*, although the beam is of small intensity, there is a distinct

broad peak in the energy spectrum. Because of the aluminium window of the scintillation probe, zero on the pulse-height scale corresponds to an incident ion energy of about 30 MeV. The highest energy recorded at 24 in. radius is about 110 MeV and it is estimated that 10 V on the pulse-height scale corresponds to about 50 MeV.

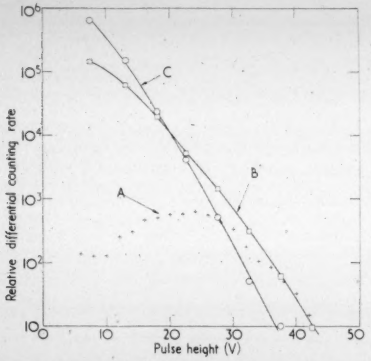


Fig. 9. Pulse-height analyses on an internal $^{14}N^{6+}$ ion beam with the scintillating crystal probe

The nature of the energy spectra conforms well to what one would expect if the ions start from the centre of the cyclotron as $^{14}N^{2+}$, undergo a preliminary third harmonic acceleration and then at some radius become stripped to $^{14}N^{6+}$ ions through collisions with gas molecules in the cyclotron tank. Since, for an orbit of a given radius, the energy of a $^{14}N^{2+}$ ion is only one-ninth of that of a $^{14}N^{6+}$ ion, the stage during which the ion undergoes acceleration in the six-charged state is the more important one in respect of the final energy achieved. If the attainment of the six-charged state occurs at a variable radius, it follows that the final energies of the ions will exhibit a continuous spread. This is indeed observed in the main region of the resonance (*B* and *C* in Figs. 4 and 9). Although it will be seen in the following section that the spectra such as *B* and *C* are affected by an increase of the gas pressure in the cyclotron tank (the spectra correspond to a low gas pressure in Fig. 10), this does not affect the qualitative argument above.

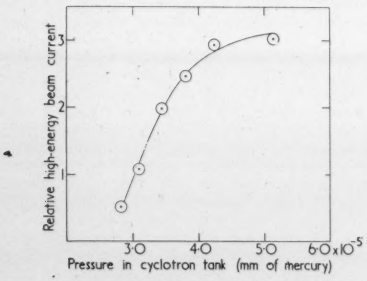


Fig. 10. Effect on a high-energy $^{14}N^{6+}$ beam of varying the pressure in the cyclotron tank

The increase in the number of high-energy ions relative to low-energy ions as the magnetic flux-density is increased can also be understood in general terms. Since at high energies the relativistic increase in mass of the ions is appreciable, the most favourable condition for the acceleration of the six-charged ions to the highest energies requires that their initial frequency of revolution at small radii be markedly higher than the frequency of the accelerating voltage on the "dees." It follows from equation (1) that the higher values of flux-density are more favourable for the acceleration of the $^{14}N^{6+}$ ions to very high energies. By contrast the initial $^{14}N^{2+}$ ions, undergoing third harmonic acceleration, suffer a negligible relativistic increase of mass. At very high values of flux-density in the $^{14}N^{6+}$ resonance (*A* in Figs. 4 and 9), the initial $^{14}N^{2+}$ ions will be able to make only relatively small excursions from the centre of the cyclotron before getting out of phase with the accelerating voltage. In this case the stripping from the doubly-charged to the six-charged state must perforce take place near the centre of the cyclotron. Thus in the extreme case *A* we do observe a tendency towards a homogeneous beam of high energy, although the total beam intensity is much less than is observed near the centre of the resonance.

7. INFLUENCE OF GAS PRESSURE IN THE CYCLOTRON TANK ON THE STRIPPING OF ELECTRONS FROM IONS

If ions such as $^{13}C^{6+}$, $^{14}N^{6+}$, $^{15}N^{6+}$ and $^{12}C^{5+}$ undergo a preliminary harmonic acceleration before being stripped to their final charge state through collisions with gas molecules in the cyclotron tank, then we would expect that the final high-energy beam should be affected by the pressure of gas in the tank. We have investigated this effect by detecting a $^{14}N^{6+}$ beam at 24 in. radius using the scintillating crystal probe. The beam was restricted to a strength which gave manageable counting rates and such a bias set on the recording system that only ions of energy greater than about one-half of the maximum (110 MeV) were recorded. The tank pressure was then raised by allowing argon gas to flow into the tank through a controllable leak. The beam current of high energy ions increased with increasing tank pressure in a striking fashion as can be seen from Fig. 10. A very similar effect was obtained when air instead of argon was admitted to the tank. When only ions of energy greater than about four-fifths of the maximum energy were recorded, a forty-fold increase of beam current was observed for the same increase of tank pressure.

If the total beam of ions of energy greater than 20 MeV at 24 in. is observed with a current-collecting probe, increase of tank pressure causes a decrease of the observed beam. Thus the increase in the number of very high energy $^{14}N^{6+}$ ions with increase of tank pressure must be accompanied by a decrease in the number of low energy $^{14}N^{6+}$ ions.

This experiment leaves no room for doubt that the basic mechanism for the acceleration of such ions as $^{14}N^{6+}$ is *via* a preliminary harmonic acceleration in the doubly-charged state. It also indicates a useful practical method of increasing the high energy components of the beam.

When using a beam of heavy ions to produce nuclear transmutations in a bombarded target, we normally run the cyclotron with a high tank pressure.

It should be pointed out that the fraction of the third

harmonic $^{14}N^{2+}$ beam, which is ultimately transformed into a first harmonic $^{14}N^{6+}$ beam of high energy, is quite small. It is difficult to determine this fraction precisely, since it depends very much on cyclotron operating conditions, but an estimate has been made by comparing the number of $^{14}N^{2+}$ ions at 20 in. radius, as measured by the thermocouple probe, with the number of $^{14}N^{6+}$ ions with energies exceeding 50 MeV at 25 in. radius, as measured by the current-collecting probe when covered with aluminium foil thick enough to stop a ^{14}N ion of 50 MeV. Under conditions favourable to the production of a high energy $^{14}N^{6+}$ beam the $^{14}N^{6+}$ beam current as specified above was not more than about 1/1000 of the $^{14}N^{2+}$ beam current.

8. THE ELECTRON-STRIPPING PROCESS

It is clear from the various experiments described that the harmonic pre-acceleration of ions is important for the ultimate production of highly-charged ions. We shall attempt in this Section to discuss the probable nature of the process in a little more detail, taking the case of $^{14}N^{6+}$ ions as an example.

The energies required to remove successive electrons from a nitrogen atom are listed in the table. The electron energies

Ionization potentials of nitrogen

Electron removed	1	2	3	4	5	6	7
Energy required in eV	14	29	47	77	97	550	650

in the cyclotron ion-source do not exceed 200 eV so that there is no possibility of the sixth electron being removed directly in the ion-source. The average number of electrons removed must certainly be much less than four, and will depend not only on the voltage applied to the arc but also on the mean free path of electrons in the gas of the arc and the possibility of recombination. However, there will certainly be good yields of $^{14}N^{2+}$ ions, the second ionization potential being only 29 V. These $^{14}N^{2+}$ ions can undergo third harmonic acceleration.

After the $^{14}N^{2+}$ ions have acquired a considerable velocity through harmonic acceleration, they may be ionized further by collision with gas atoms inside the cyclotron tank. To appreciate the order of magnitude of the energy required by the $^{14}N^{2+}$ ion, if in a collision with an atom of thermal energy there is to be a good probability of removing the sixth electron from the $^{14}N^{2+}$ ion, we can, as a first approximation, picture the struck atom simply as a cloud of stationary electrons which the nitrogen ion encounters. If v is the relative velocity of the $^{14}N^{2+}$ ion and a stationary electron which makes a head-on collision with the sixth electron of the $^{14}N^{2+}$ ion, m the electron mass and I the binding energy of the sixth electron in nitrogen, then

$$\tfrac{1}{2}mv^2 = I \qquad (3)$$

If M is the mass of the ^{14}N ion, then its kinetic energy E will be given by

$$E = I(M/m) \qquad (4)$$

with $I = 550$ eV, $E = 14{\cdot}3$ MeV. Since, however, the electrons in the struck atom are in fact in motion with respect to their nucleus and, further, there is a continuous distribution in the energy transferred in an electron-electron collision, there is no sharp threshold in the $^{14}N^{2+}$ energy required. There is, rather, a continuously increasing probability of removing the sixth electron as the energy of the nitrogen ion is increased upwards from zero, the probability approaching unity only at energies above 14·3 MeV. In the cyclotron the energy of a third harmonic $^{14}N^{2+}$ ion at a radius of 24 in. is 12 MeV. An estimate based on the calculations by Brunings, Knipp and Teller[15] of the average charges carried by heavy ions passing through matter indicates that, at 12 MeV, ^{14}N ions passing through matter would carry a mean charge of between 5+ and 6+.

Observations on the strength of the $^{14}N^{2+}$ beam in the cyclotron are consistent with the collision cross-section for further ionization of the $^{14}N^{2+}$ ions being of the order of magnitude of the geometric cross-sections of the gas atoms in the cyclotron tank. It must be pointed out, however, that a single collision of a fast $^{14}N^{2+}$ ion with a gas atom seems unlikely to remove four electrons simultaneously from the nitrogen ion and leave it directly in the six-charged state in which first harmonic acceleration is possible. Rather it seems probable that nitrogen ions which have undergone third harmonic acceleration will, after making a collision causing further ionization, circulate in the tank with a state of charge intermediate between 2+ and 6+. Such ions, not being resonant, will on the average neither gain nor lose energy. They will make further collisions, some of which will lead to the loss of electrons and some to the capture of electrons. A state of equilibrium will be reached in which a certain fraction of the non-resonant ions will be six-charged and of these six-charged ions those with appropriate orbits and phases with respect to the radio-frequency accelerating voltage will be accepted into first harmonic acceleration.

One would expect, on the basis of the foregoing discussion, that the high energy $^{14}N^{6+}$ beam current would be very small compared with the third harmonic $^{14}N^{2+}$ beam current since (*a*) $^{14}N^{2+}$ ions which make their first ionizing collision at a small radius will not have a high probability of ultimately becoming $^{14}N^{6+}$ ions and (*b*) only those $^{14}N^{6+}$ ions which fulfil particular kinematic conditions can be accepted into first harmonic acceleration. The low efficiency of the whole process is confirmed by the measurement described at the end of Section 7.

Most of the useful ionizing collisions will take place at large radii. This is not only because the greater part of the $^{14}N^{2+}$ path lies at large radii on account of the spiral nature of this path but also because the energy of the $^{14}N^{2+}$ ions increases as the square of the radius. Now although the effect will be reduced if $^{14}N^{2+}$ ions do not in general strip directly to $^{14}N^{6+}$ ions, $^{14}N^{2+}$ ions which make their first ionizing collision at a large radius will tend to give rise to $^{14}N^{6+}$ ions of low final energy. This is because there is in this case relatively little radius available for the regime of first harmonic acceleration of $^{14}N^{6+}$ and it is this regime which is important from the point of view of the final energy achieved, as has been pointed out in Section 6(*b*). The prominence of low energy ions in the energy distributions obtained near the centre of the magnetic resonance curves can thus be understood. The effect of increased gas pressure in increasing the number of high energy ions and decreasing the number of low energy ions (Section 7) can also be understood since more ionizing collisions take place at small radii when the pressure is raised. Of course, the effectiveness of the individual collisions at small radii is not increased, but the number of such collisions is increased.

The relatively very small efficiency of fifth and seventh harmonic pre-acceleration when compared with third harmonic pre-acceleration has been mentioned in Section 5.

This can be understood for several reasons. First, at a given radius the velocity of a fifth harmonic ion is only 3/5, and that of a seventh harmonic ion only 3/7, of the velocity of a third harmonic ion. Consequently the probability of useful further ionization is reduced. Secondly, the initial state of charge of $^{12}C^{1+}$ and $^{16}O^{1+}$ is so low that the first collision which these ions make near the ion source seems likely to lead to the loss of a further electron, thus destroying the fifth or seventh harmonic relationship. Lastly, especially at small radii, the low angular velocity of fifth and seventh harmonic ions leads to large changes in the phase of the radio-frequency accelerating voltage occurring while the ions cross the gap between the "dees" with consequent inefficiency of acceleration.

A final remark should be made concerning the electron-stripping process. It might at first sight be thought that it would be better for the purpose of producing large beams of $^{14}N^{6+}$ to use a very high voltage ion-source, such as a spark discharge, to produce $^{14}N^{6+}$ ions directly instead of utilizing the third harmonic pre-acceleration of $^{14}N^{2+}$ ions. However, even if such six-charged ions were produced in the ion-source they would capture electrons in subsequent collisions in or near the source while still at low energies and consequently would not survive in appreciable numbers. The small peak X in Fig. 6 ($^{12}C^{5+}$ resonance) may possibly be due to a small number of $^{12}C^{5+}$ ions produced near the ion source, such direct acceleration of $^{12}C^{5+}$ ions being observable only because of the very poor efficiency of fifth harmonic pre-acceleration.

9. Some General Remarks on Beam Intensity

We shall mention here some miscellaneous points which have not so far been touched upon.

(*a*) *Effect of the amplitude of the radio-frequency accelerating voltage.* In order to produce the best beams of high energy, highly-charged ions, it is always found to be an advantage to use as high an accelerating voltage as possible. This is the general experience in accelerating any ion to high energies in a fixed-frequency cyclotron. The reason is that the relativistic increase in mass of an ion, combined with the fall-off in the flux density between the centre and outside of the cyclotron (necessary to provide vertical focusing), reduces the frequency of revolution of the ion as its energy is increased [equation (1)]. This makes it desirable to get an ion up to full energy before it gets out of phase with the radio-frequency accelerating voltage. With heavy ions such as those of ^{14}N which make a transition from third harmonic (N^{2+}) to first harmonic (N^{6+}) acceleration, there is likely to be an additional reason for making a large accelerating voltage desirable. It is best that a $^{14}N^{6+}$ ion which has been accepted into first harmonic acceleration be accelerated as rapidly as possible so that not only will the number of its subsequent collisions be reduced but also so that any such collisions will be less likely to lead to a reduced degree of ionization.

(*b*) *Ion source conditions.* In accelerating carbon, oxygen and nitrogen ions, we have found with the usual hooded capillary ion source that the highest yields of doubly-charged ions are obtained under conditions similar to those which yield the highest beams of helium ions. The chemical nature of the gas supplied to the ion source, however, appears to be of importance when producing $^{16}O^{2+}$ ions. Nitrous oxide gives about ten times the yield of $^{16}O^{2+}$ ions which is obtained with either oxygen gas or a mixture of oxygen and carbon tetrafluoride. On the other hand when $^{14}N^{2+}$ ions are being produced, nitrogen gas, nitrous oxide, or a mixture of nitrogen and carbon tetrafluoride appear about equally efficient. These effects have not been fully explored with various gases, but it seems that the nature of the valency binding in nitrous oxide has an effect on the yield of O^{2+} ions.

(*c*) *Maximum beam strengths obtained.* With highly-charged ions we shall specify the high-energy beam current as that of all particles with energies exceeding 50 MeV at 25 in. radius. With $^{14}N^{6+}$ we have observed beam currents of up to about 5×10^{-9} A, the corresponding third harmonic beam at 20 in. radius being several μA. With $^{15}N^{6+}$, using nitrogen gas of normal isotopic constitution in the ion source, the beam currents are only about 0·3% of this. With $^{13}C^{6+}$, the use of carbon dioxide gas of normal isotopic constitution in the ion source has led to maximum beam currents of about 10^{-10} A. With $^{16}O^{6+}$, using nitrous oxide in the ion source, beam currents of about 10^{-10} A have been obtained. No attempt has been made to readjust the magnetic "shims" of the cyclotron, normally used for the acceleration of deuterons, to obtain optimum conditions for the acceleration of these heavy ions.

Acknowledgements

We should like to express our appreciation to Professors P. B. Moon and W. E. Burcham for the interest they have shown in this work and to members of the cyclotron group of the University of Birmingham for their co-operation.

References

(1) Freier, P., Lofgren, E. J., Ney, E. P., Oppenheimer, F., Bradt, H. L., and Peters, B. *Phys. Rev.*, **74,** p. 213 (1948).

(2) Alvarez, L. W. *Phys. Rev.*, **58,** p. 192 (A) (1940).

(3) Miller, J. F., Hamilton, J. G., Putnam, T. M., Raymond, H. R., and Rossi, G. B. *Phys. Rev.*, **80,** p. 486 (L) (1950).

(4) Ghiorso, A., Thompson, S. G., Street, K. and Seaborg, G. T. *Phys. Rev.*, **81,** p. 154 (L) (1951).

(5) Miller, J. F. *Phys. Rev.*, **83,** p. 1261 (L) (1951).

(6) Chou, C. N., Fry, W. F., and Lord, J. J. *Phys. Rev.*, **87,** p. 671 (L) (1952).

(7) Walker, D., and Fremlin, J. H. *Nature* [*London*] **171,** p. 189 (1953).

(8) Wyly, L. D., and Zucker, A. *Phys. Rev.*, **89,** p. 524 (L) (1953).

(9) Chackett, K. F., Fremlin, J. H., and Walker, D. *Proc. Phys. Soc.* [*London*] *A*, **66,** p. 495 (1953).

(10) Breit, G., Hull, M. H., and Gluckstern, R. L. *Phys. Rev.*, **87,** p. 74 (1952).

(11) Giles, P. C., and Barkas, W. H. *Phys. Rev.*, **85,** p. 756 (A) (1952).

(12) "University of Birmingham Cyclotron." *Nature* [*London*] **169,** p. 476 (1952).

(13) Condit, R. *Phys. Rev.*, **62,** p. 301 (A) (1942).

(14) Link, W. T., and Walker, D. *Proc. Phys. Soc.* [*London*] *A*, **66,** p. 767 (1953).

(15) Brunings, J. H. M., Knipp, J. K. and Teller, E. *Phys. Rev.*, **60,** 657 (1941).

Summarized proceedings of a conference on electron microscopy—London, November 1953

The Annual Conference of the Electron Microscopy Group of The Institute of Physics was held in the Physics Department of Birkbeck College, University of London, on 10 and 11 November, 1953.

In common with previous conferences of the group, papers were read covering the electron microscope and a wide range of its applications and the proceedings are summarized in this report.

ELECTRON OPTICS

MR. M. E. HAINE and MR. T. MULVEY (Associated Electrical Industries Ltd., Aldermaston) read the first paper of the conference on the correction of astigmatism. They analysed the edge diffraction test of Hillier and Ramberg[1] and described the design and use of a four-pole cylindrical stigmator lens. It was demonstrated that the edge diffraction test was theoretically sensitive enough to allow the reduction of astigmatism to a value where it becomes negligible in comparison with spherical aberration and diffraction. On the other hand, the diffraction fringe width, which must be observed in the course of the test, approaches a value equal to the resolving power, as limited by spherical aberration and diffraction, as the astigmatism becomes negligible. It is therefore a necessary condition for the application of the test, that no other factors shall limit resolution to a worse value than that set by spherical aberration and diffraction. Conversely, the successful application of the test demonstrates that no such extra limitation to resolution exists.

The astigmatism corrector used was of the four-pole electrostatic type with mechanical rotation. Correction requires only a few volts. Experiments have shown that a resolving power of better than 10 Å is regularly obtained by the use of this corrector. Edge diffraction fringe widths (peak ,to minimum) of 4–5 Å have been recorded. Particles of aggregated gold-palladium on a carbon film, having an average size of 10 Å, are completely separated at a spacing of 15–20 Å indicating a resolving power of 8–10 Å. The necessity for recording images at magnifications sufficiently high to avoid the effects of statistical fluctuations in the electron beam was stressed. The authors then suggested that there appears to be no reason why future commercial instruments should not have a resolving power very close to 10 Å. In the ensuing discussion Mr. Haine was questioned about his alinement procedure and he pointed out that the "voltage" centration and "current" centration procedures described by Hillier and Ramberg give the same axis if the lens design is adequate and iron saturation effects are avoided.

CHEMICAL APPLICATIONS

MR. E. CRAMPSEY, DR. M. GORDON and MR. J. W. SHARPE (Royal Technical College, Glasgow) presented a paper describing an electron microscopy study of the hydrochlorination of natural and synthetic polyisoprene. Latex samples were examined shadowed and unshadowed and the discrepancies in size described by Kern and Kern,[2] and Cosslett[3] between the two sets of results were noted. Micrographs of particles before hydrochlorination appeared rubbery and flattened, while afterwards they appeared glassy and stiff. The method was put forward as being a possible tool in vulcanization studies.

MISS K. LITTLE (Medical Research Council Radiobiological Unit, Harwell) described two ways in which the electron beam can act on organic high polymers. It can act as a source of ionizing radiation, and also as a source of thermal energy. In the first case the effects have been indistinguishable from those produced by other radiation sources. They are primarily chemical, and depend upon the chemical nature of the polymer. Some polymers such as cellulose are mainly degraded, while with others such as proteins and nucleic acids a rearrangement of bonds is possible. Changes in samples bombarded with electrons from a van de Graaf generator were followed using X-ray diffraction methods. Changes in solubility and melting point sometimes occurred, and when the system was such that the beam charred the specimen, changes in its appearance were noted. In the discussion DR. G. EAVES (University of Leeds) pointed out that when he subjected sections to a high beam intensity a high proportion of the plastic embedding evaporated, and he contended that this indicated the thermal effect to be the more important in the case of methacrylates.

MISS J. M. PESKETT (Mechanical Engineering Research Laboratory, East Kilbride) described a method used for specimen preparation of the soap fibres present in alkali-base greases. The oil was vaporized in a vacuum to leave the soap fibres unchanged in size, shape and arrangements for viewing in the electron microscope. The specimens were used in research into the function of the structure of a grease in the lubrication of roller bearings.

APPLICATIONS OF THE GRAZING INCIDENCE MICROSCOPE

Dr. Menter and his colleagues (Research Laboratory of the Physics and Chemistry of Surfaces, University of Cambridge) presented three papers describing research done with the aid of the converted Metropolitan-Vickers type E.M.3 machine which was described at last year's conference. In the first of these papers, MR. J. A. CHAPMAN and DR. J. W. MENTER showed that the reflexion technique could be used to study the shape and surface structure of synthetic and other fibres provided that the specimen was suitably mounted and the intensity of the illuminating electron beam was kept to a minimum. Micrographs of viscose rayon, cellulose acetate, nylon, terylene, orlon and cotton taken by the method were shown and the surface structure of drawn

and undrawn nylon model filaments was examined in some detail. The maximum resolution obtained, so far, was of the order of 1000 Å, but this should be improved. The chief advantage of the method is in the very high depth of field. This overcomes the main drawback of the light microscope by presenting a picture of the fibre as a whole. In the discussion MR. H. W. EMERTON (British Paper and Board Research Association, London) referred to his collaboration with Mr. Chapman and Dr. Menter in the preliminary examination of plant fibres in the reflexion microscope. He then outlined the development of this work with Amboss and Watts at the Research Association's laboratory and showed slides of plant fibres taken with the reflexion microscope (Fig. 3).

DR. J. F. HALLIDAY (Associated Electrical Industries Ltd., Aldermaston) described how he had used reflexion electron microscopy in examining surface finish; he found that the method was particularly suitable for the study of surfaces prepared by typical engineering methods such as grinding, lapping, abrading and diamond turning (Fig. 2). Although a single reflexion electron micrograph can be fully interpreted more information can usually be obtained by a comparison of micrographs either (1) by examining various surface finishes on one or more metals at the same angles of illumination and viewing, or (2) by examining a particular area at a constant angle of viewing and at several decreasing angles of illumination. Several slides of abraded mild steel and diamond turned copper illustrated method (2). On the abraded mild steel, the majority of the surface asperities sloped less steeply than 2° in the direction of abrasion, and on the diamond turned copper many of the asperities had slopes of less than ½°; such details cannot be seen unless very small angles of illumination are used. For this reason, the reflexion method is considered to be superior to the replica method for the study of surface finish since the latter technique has not yet been adapted for shadowing at sufficiently small angles.

The reflexion method is also very suitable for the study of the surface damage resulting from sliding. After sliding, torn particles may remain adhering only loosely, and such particles might easily be removed when taking a replica. The reflexion method does not possess this disadvantage; in suitable circumstances it can show directly not only the adhering particle but also the cavity from which it was torn. Fig. 1 shows an example of this and it will be noted from its shadow that the upstanding flake has the same outline as the adjacent depression in the surface.

MR. S. A. WATSON (British Non-Ferrous Metals Research Association, London) and DR. J. W. MENTER described how the reflexion technique has been used to study the character of chemical and electrodeposits of zinc on aluminium. The structure of the zinc layer has an important bearing on the adhesion of subsequent layers of other metal plated on to the surface. The observations have shown that under conditions giving the best adhesion, the zinc layer is homogeneous both in particle size and in the distribution of particles on the surface (Fig. 4).

The reflexion electron microscope overcomes the main difficulty with the light microscope in which it is almost impossible to distinguish the electrodeposited nuclei from the etch structure of the underlying metal. It was suggested that the method may find widespread application in this field where it is desired to resolve electrodeposited particles which are often just beyond the resolution limit of the light microscope.

MR. M. SEAL and DR. J. W. MENTER described the examination of polished surfaces of diamonds by the reflexion technique. The stones examined had all been polished in one direction only and, in addition to the grooves formed by the abrasive, they showed a number of other lines different in character and direction. These lines follow the traces of (111) planes of greatest resolved shear stress (Schmid criterion) and not of greatest normal breaking stress. If a diamond face is repolished in a different direction the lines change direction accordingly. Contamination films formed on the surface of the specimens in the reflexion microscope have been stripped and used in the transmission microscope as very low angle shadowed replicas for estimating the change in surface height at the lines. This may, in some cases, be as small as 20 Å and the horizontal separation of the lines is about ½ μ. The lines are almost certainly due to crystallographic slip occurring during polishing. The possibility of plastic deformation in diamond has an important bearing on theories of the mechanism of the polishing of this material.

DR. S. P. NEWBERRY (General Electric Co., Schenectady, N.Y.) and DR. W. C. NIXON (Cavendish Laboratory, Cambridge) read a short paper on some recent results of X-ray microscopy in America.

A comparison of electrostatic and electromagnetic lenses for the X-ray shadow microscope shows that the intensity or resolution is up to twenty times better for the magnetic case. X-ray micrographs obtained with magnetic lenses were shown including test specimens of 1500 mesh per in. silver grid, 3 μ bar width, demonstrating approximately 1 μ resolution. Among the metallurgical results an aluminium-tin alloy, 5% tin, 0·020 in. thick, showed the deposition of tin along the aluminium grain boundaries. This was shown more clearly in a stereographic X-ray micrograph of this specimen, taken on Polaroid-Land film and developed for 60 sec in the camera.

A test grid specimen showed that Dr. Baez' X-ray shadow microscope at the University of Redlands, California, has obtained a resolution of 1 μ.

REPORT OF GERMAN CONFERENCE

DR. V. E. COSSLETT and MR. M. E. HAINE opened the first session of the second day of the Conference with a brief report of the German Conference held at Innsbruck on 16–19 September, 1953. An abstract will not be attempted here as full reports are available.[4]

REPLICA TECHNIQUES

MR. D. E. BRADLEY (Associated Electrical Industries Ltd., Aldermaston) read the first paper of a session on replica techniques. He gave a comprehensive description of a new technique in the preparation of specimen supporting membranes and replicas. Carbon is evaporated by passing a current of about 30 A through two pointed carbon rods held lightly together in a vacuum and a film of carbon is deposited in an amorphous form suitable for specimen support films and for forming replicas. Its advantages are that it is highly transparent to electrons, and very strong and stable in the electron beam. Very thin specimen support films can be made by using the resin Bedacryl by I.C.I. Ltd. as a substrate and washing it away with an ether-acetone mixture after mounting the combined carbon film and substrate on a grid. Collosion and formvar have also been used as substrates.

Fig. 1. Centre of track made on abraded copper by $\frac{3}{8}$ in. diameter steel ball, load 200 g. $\theta_1 = 1°$. $\theta_2 = 4°$. Slide direction—from top to bottom of print. Note undulations in track and complete pictures of holes, etc.
(J. F. Halliday, A.E.I.)

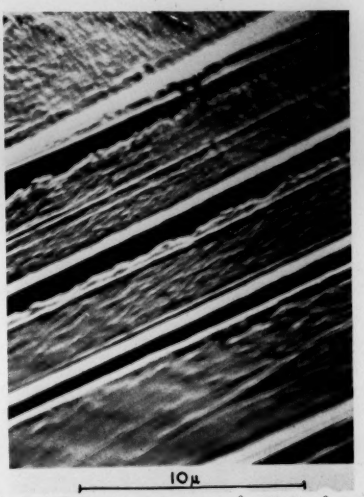


Fig. 2. Diamond turned copper. $\theta_1 \simeq 0{\cdot}2°$. $\theta_2 = 4°$. Lines cast no shadows until $\theta_1 < 1°$. Asperities between lines become visible when $\theta_1 < 0{\cdot}3°$. Diamond motion from top right to bottom left of print.
(J. F. Halliday, A.E.I.)

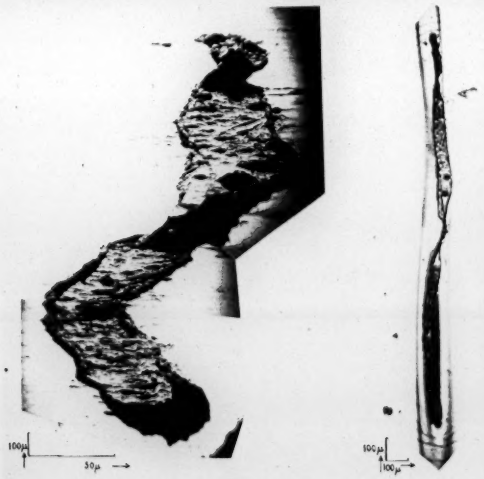


Fig. 3. Wood fibre. *Left*—reflexion electron micrograph. *Right*—light micrograph.
(H. W. Emerton, British Paper and Board Research Association)

Fig. 4. Electrodeposit of zinc on etched aluminium surface with current density of 4·83 mA/cm^2, vertical magnification = 6250 (beam travelling from left to right).
(J. W. Menter and S. A. Watson, Cambridge)

Figs. 5 and 6. Evaporated carbon replicas of pearlite structure in a heavily etched specimen. ×10 000
(D. E. Bradley, A.E.I.)

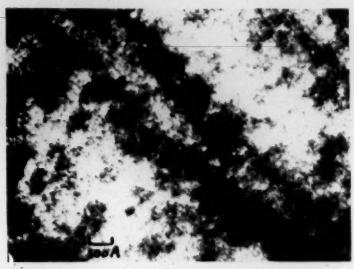


Fig. 7. Intracellular double membrane from a section of C3H mouse breast tumour. ×500 000
(G. Eaves, Leeds)

Fig. 8. Section of vaccinia virus in chorio allantoic membrane. ×135 000.
(T. H. Flewett and G. Eaves, Leeds)

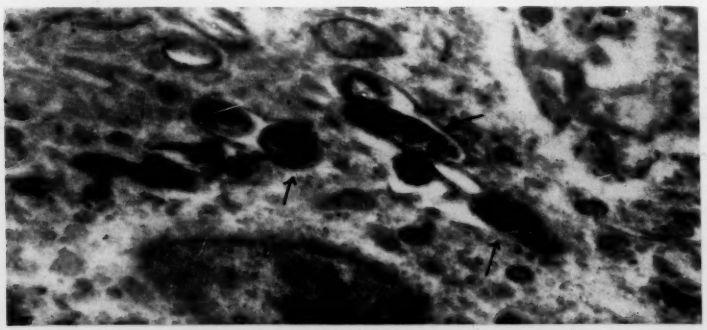


Fig. 9. Section of mouse being infected with bovine tubercle bacillus. Bacteria are arrowed. ×28 500.
(E. M. Brieger and A. M. Glauert, Cambridge)

For normal specimen supports of about 50 Å thickness, carbon may be evaporated on to glycerol, or plain glass cleaned with detergent. The film is then floated on a water surface and picked up on grids.

Two replica techniques have been developed; one for the examination of specimens which can be dissolved away from the final replica, and a two-stage method for the production of positive replicas from metals and similar specimens.

The single-stage method has been applied to the examination of photographic grains. An aqueous suspension of the particles is dried on to a microscope slide, carbon is evaporated normally on to the surface and the halide then removed by solution in sodium thiosulphate. The replicas clearly delineate the crystal shape and have a resolution of better than 100 Å.

The behaviour of carbon when evaporated on to a surface was discussed. A slide of a replica of polystyrene latex particles, made by evaporating the carbon on at $\tan^{-1}\frac{1}{2}$ showed sharp shadows, but some detail was visible in the shadowed area. This was because a proportion of the carbon appears to travel in a straight line from the source and the remainder forms a thin film over the entire surface due to flowing, gas scattering, or reflexion from adjacent pieces of apparatus.

The two-stage technique described consists of dry-stripping a thick formvar replica backed with Bedacryl from a prepared metal surface by means of cellulose adhesive tape. Carbon is evaporated on to the formvar which is then floated free by immersing the tape in acetone which dissolves the Bedacryl. The composite film is then mounted on a grid, and the formvar removed by washing with chloroform. Micrograms of unshadowed pearlite structure in annealed carbon steel were shown illustrating the difficulty in interpreting results, but after shadowing the nature of the surface was clarified (see Figs. 5 and 6).

The main advantage of these methods over existing silica techniques is their simplicity. The replicas are suitable for stereoscopy. They give a resolution of better than 50 Å.

Mr. J. F. Nankivell (Cavendish Laboratory, Cambridge) read a short note followed by another by Mr. E. Smith and Dr. J. Nutting (Metallurgy Laboratory, Cambridge) illustrating some applications of Mr. Bradley's technique. They showed micrographs of a number of carbon replicas many of which demonstrated a considerably higher resolving power than has hitherto been obtained using plastic replicas. Dr. C. E. Challice (Wright-Fleming Institute, London) described how he modified Mr. Bradley's technique in making supports for biological objects. He used a thin Bedacryl substrate, deposited carbon, and then mounted the biological specimen. The thin Bedacryl was largely vaporized in the electron beam.

Unexpected aspects of the polystyrene-silica replica process, which are demonstrated when the method is applied to the highly curved surface of a textile fibre, were described by Mr. S. F. Ward and Dr. D. G. Drummond (Shirley Institute, Manchester). When silica is evaporated on to the polystyrene it covers the entire surface, but remains thinner in the lee of a projection than elsewhere. Asperities on those parts of the surface which are steeply inclined to the silica stream therefore cast shadows, which become visible in the electron microscope because the silica replica tends to flatten out when the polystyrene is removed. This self-shadowing facilitates the detection and interpretation of fine details of the fibre surface.

Mr. G. R. Booker (Metropolitan-Vickers Electrical Co., Manchester) described a simple method for examining selected areas of surfaces using wet-stripped replicas. The selected area of the surface is identified by fine scratches in the immediate vicinity. The corresponding portion of the replica is mounted over a 0·015 in. diameter hole in a copper disk and viewed in an electron microscope operating with a specimen screening aperture. The method has several advantages over methods hitherto used, for example it has extremely wide application as it does not depend on surface structure for identification; the positioning of the replica on the copper disk is rapid and can be correctly performed each time, and an uninterrupted area 0·015 in. in diameter can be observed.

Miss M. K. B. Day (British Aluminium Co., Ltd.) described how the anodic replica technique has proved to be particularly suitable for studying the structures of cast aluminium-manganese alloys as it reveals both the coarse (up to 4·0 μ) and very fine (about 0·1 μ) particles of constituent with the same degree of contrast. The specimen was first polished so as to avoid relief effects; then polished mechanically on diamond dust, and finally given a very short electropolish to clean up the aluminium background. At first great difficulty was experienced in stripping the anodic replicas off these alloys, but this was overcome by saturating the surface layer of the specimen with hydrogen prior to anodizing—initially the sample was made the cathode in the anodizing bath for 1 min before the connexions are reversed. The films were then stripped off on mercuric chloride solution.

Dr. G. Eaves (University of Leeds) opened a short discussion on the cutting of thin sections such that the full resolving power of the electron microscope was used. To attain a resolution of 10 Å he suggested that the section must be 100 Å in thickness or thinner, and showed several electron micrographs of sections which demonstrated the good resolution which he had attained. One picture, enlarged greatly from a fairly low magnification, showing an apparent grain which photographic experts had said was not photographic grain (Fig. 7) was discussed at some length by Mr. Haine (Associated Electrical Industries Ltd., Aldermaston) who reported some interesting work on the blackening of photographic emulsions by electrons. He said that the graininess in the photographic emulsion for electron exposure is considerably greater than for light exposure. The reason lies in the relatively greater efficiency of electrons in blackening the plate, and in the random fluctuations in the electron beam. Thus a 10 Å diameter area magnified $\times 10\,000$ requires only about 250 electrons to blacken it to a density of 1. Statistical fluctuations will cause a variation in this number of $1/\sqrt{250}$, or 8·5%. With a γ-value of 3 the contrast variation will be 25%. Mr. Haine said that the experiments carried out at Aldermaston confirmed the general accuracy of this argument. They found that to avoid random graininess it is necessary to photograph a 10 Å resolution image at a magnification of at least $\times 50\,000$. The electron grain effect bears no relation to photographic grain, but is dependent on the speed of the emulsion.

Miss K. Little (Medical Research Council Radiobiological Unit, Harwell) questioned the manner in which the methacrylate embedding of sections was assumed to be evaporated when the electron beam was applied as she had found (as described in her paper read on the previous day) that changes of structure in plastics could be induced in this way. Several ultra-microtomists asserted that they had never observed any structure develop in the methacrylate embedding.

MR. R. C. VALENTINE and DR. J. R. G. BRADFIELD (University of Cambridge) described studies on the survival of bacteria (*paracolon bacilli*) suspended in a small drop of distilled water on an electron microscope specimen grid. Survival is zero after 10 min, the bacteria apparently being killed when they break through and lie on the surface of the drop. There is no significant fall in survival when the bacteria stand for 10 min in a flask of distilled water. If the drop is allowed to evaporate, there are no survivors except in those areas from which the drop has retracted before drying; these areas contain many live, dried bacteria. From these observations they developed a simple draining method for preparing bacteria for electron microscopy, which reduces cell damage and avoids centrifugation. They finally suggested that, in addition to the surface killing, drying bacteria may be frozen to death; as a result of the greatly increased rate of evaporation given by kinetic theory for drops of diameter less than 10 μ.

In another paper DR. J. R. G. BRADFIELD and MR. R. C. VALENTINE described preliminary observations on 0·035 μ sections of *Staphylococcus aureus* and a *paracolon bacillus*. Both show well-defined cell walls, both lack mitochondria and both sometimes show the peripheral cytoplasmic vesicles associated with cell division. Main interest, however, centres in the nuclei. Actively growing staphylococci (0·5 μ diameter) possess a well-defined spheroidal nucleus (0·2 μ diameter) containing bodies looking very much like chromosomes and nucleoli. In osmium-fixed cells the nucleus is much less dense to electrons than the cytoplasm. By means of a modified Feulgen stain, which deposits silver granules in sites containing desoxyribonucleic acid (DNA), it is possible to show in the electron microscope that DNA is localized in the bacterial nuclei, which therefore have both the morphological and the chemical characteristics of animal and plant nuclei. In a dividing staphylococcus the nucleus elongates and nips into two; there is no evidence of true mitosis. Actively growing paracolon bacilli (2 μ by 0·6 μ) possess irregular ellipsoidal nuclei approximately 1·5 μ by 0·4 μ. Nuclear : cytoplasmic ratios are 1 : 10 in the young cocci and 1 : 2 in the young bacilli. In the older resting bacilli the nucleus shrinks, becoming extensively branched, for reasons unknown. It is now clear that in *paracolon bacilli* (and probably in *E. coli*) the nuclei, in spite of their DNA content, give rise to the less dense, central areas, of variable shape, which have long been familiar in electron micrographs of the whole bacteria.

The internal structure of vaccinia elementary bodies as seen in thin sections of infected egg membranes was described by DRS. T. H. FLEWETT and G. EAVES (University of Leeds). The virus particles consisted of a dense cortex, which showed concentric lumination, surrounding a less dense central structure about 60 × 100 mμ across as seen in section. The advantage of high resolution electron micrographs of sections, which enabled individual virus particles to be distinguished from small mitochondria and other cell inclusions, was pointed out (Fig. 8).

DRS. E. M. BRIEGER and AUDREY M. GLAUERT (Strangeways Research Laboratory, Cambridge) gave an introductory outline of the development of avian and bovine tubercle bacilli on artificial media. Electron micrograms of intact bacilli were shown to illustrate the reproductive cycle. The technique employed to obtain ultra-thin sections of infected tissues from rabbits and mice was then described. After fixation with osmic acid and embedding in *n*-butyl methacrylate, the sections were cut on Porter's microtome with a glass knife. Electron micrograms showed the form and location of tubercle bacilli in the lung and spleen of a heavily infected mouse. They were accompanied by other inclusions, some of which were apparently vacuolized while others were dense and granular. Sections of tissues from a rabbit killed at an earlier stage of the disease did not contain any bacilli but dense intracellular inclusions were seen similar to those accompanying bacilli in mice. These were presumed to be the infective agents. It was concluded that the method is of considerable value for the study of the form, location and development of bacilli in infected tissues (Fig. 9).

MRS. B. E. WILLIAMS (National Physical Laboratory, Teddington) informed the Conference that the N.P.L. had found a method of producing gelatin copies of diffraction gratings at negligible cost. These could be used many times for producing replicas suitable for the electron microscope, and the N.P.L. were prepared to supply them, at a nominal charge (to cover postage and packing), to electron microscopists who wished to use them for magnification calibration.

C. E. CHALLICE

REFERENCES

(1) HILLIER, J., and RAMBERG, E. G. *J. Appl. Phys.*, **18**, p. 48 (1947).

(2) KERN, S. F., and KERN, R. A. *J. Appl. Phys.*, **21**, p. 705 (1950).

(3) COSSLETT, V. E. *Proc. Electron Microscopy Conf.* Delft (1950).

(4) *German Electron Microscopy Conference.* Reports in:
Physik Verhand, **4**, pp. 83–124 (Abstracts).
Z. Wiss. Mikr. To be published.
Optik. To be published.

Some measurements of the relative dielectric strength of gases

By N. R. McCormick, B.Eng., Ph.D., and J. D. Craggs, M.Sc., Ph.D., F.Inst.P., Department of Electrical Engineering, The University of Liverpool

[*Paper first received* 12 *June, and in final form* 1 *July*, 1953]

Measurements of the breakdown voltages of small sphere gaps in various gases have been made, partly to check existing results but also to provide new data. Various electro-negative gases, such as SF_6, Cl_2, etc., were studied, taking nitrogen as a standard gas, and the data are tabulated. A brief discussion of these results is given.

The technical importance of certain gases, for example SF_6 and CCl_2F_2, as insulators is increasing and their behaviour is also of basic interest. Thus, it has been tacitly and, we believe, rightly assumed by several authors[1] that the formation of negative ions during the electrical breakdown process in such gases is responsible for their relatively high dielectric strength (i.e. about 2–3 times greater) as compared, for example, with that of nitrogen. This general assumption is seldom taken further and, indeed, the work of Penning[2] on oxygen using values of Townsend's α coefficient due to Masch was the first to show specifically how negative ion formation can quantitatively effect the latter (α) coefficient, which is the most important single quantity governing electrical breakdown in gases. It should be mentioned that the values of Townsend's γ coefficients in the strongly electro-negative gases do not yet appear to have been measured, although they are being studied in this laboratory. Breakdown mechanisms are discussed elsewhere by Llewellyn Jones and his collaborators,[3,4] who have measured the Townsend coefficients in air and nitrogen up to the threshold of breakdown. Loeb[5] has given a general discussion of secondary mechanisms in gaseous breakdown.

The present series of measurements was undertaken to provide preliminary data on relative dielectric strengths, with an accuracy of about $\pm 5\%$. Many such measurements have been made and references are given elsewhere[6,7] and in Table 2, which gives a selection of experimental data. Certain discrepancies appear, for example with respect to the breakdown voltages for Cl_2, and some of these were investigated. Also it was desirable, in connexion with other experiments being made with the electron attaching gases, including the investigation of dissociation patterns in a mass spectrometer, to obtain new data, for example, on the series CF_3X where X = F, Cl, Br or I.

The measurements were made with small gap spacings in sphere-sphere gaps, where the ratio of sphere diameter to gap spacing was about 10–20 : 1. This gives a close approach to uniform field conditions.[8,9] Pressures of about 1 atm were generally used but this was sometimes inconveniently high, as in the case of C_2N_2, and data at lower pressures were then obtained.

APPARATUS

A simple spark chamber was built with a $3\frac{1}{4}$ in. internal diameter glass cylinder sealed to Fe–Ni–Co cylinders flared out to form flanges. The latter may be brazed into metal end plates or, as in the present experiments, clamped to them with rubber gaskets. These plates carry the (brass) spheres, one of which is operated by means of a flexible bellows joint from outside the chamber. The impurity due to residual air was, for 1 atm of the test gas, about 1 in 10000, which for the present experiments was considered adequate. It is perhaps appropriate to point out here that many of the complex molecular gases, for example CCl_2F_2, dissociate readily in a discharge even when the current is limited, as in the present case, to $\simeq 1$ mA, and for this reason it is difficult to reach the precision obtainable with pure elementary gases or with free air. The effects of permanent dissociation on the measurement of breakdown voltages can readily be detected by repeated sparking and was not appreciable in the present work. The zero reading for the sphere gap could be found by bringing the spheres into contact, using an external micrometer screw fixed to one end plate of the chamber, then the gap spacing could be set as required. The breakdown voltages (d.c.) were measured with a calibrated resistance potential divider and an electrostatic voltmeter reading to about $0{\cdot}5\%$ of full scale, also suitably calibrated. The procedure was to increase the voltage applied to the gap slowly by manual control of an electronically stabilized high voltage source, using some $0{\cdot}5$ mg of radioactive cobalt placed near the test gap to provide irradiation, until breakdown was obtained, as indicated by a spark and the collapse of the voltmeter reading.

The consistency of this method may be judged by the following breakdown fields (arbitrary units) for nitrogen, for 2 mm sphere-sphere spacing and about 815 mm of mercury pressure, over a period of some months with frequent dismantling of the chamber for cleaning: 123, 120, 120, 123, 123, 120, 121, 122, 122, i.e. 121 ± 2. This consistency is better than that claimed above for the complex gases. The procedure was to take a run with N_2 in the chamber for various gap spacings, then to introduce the test gas, make a similar series of measurements, finally making another run with N_2.

In most cases (CCl_3F, CCl_2F_2, $CClF_3$, CF_4, $CBrF_3$, CIF_3, BF_3, SF_6 and C_2N_2) mass spectrometer studies of the gases have been made, primarily to establish dissociation patterns involving negative and positive ions. The measurements have also been of value in assessing the impurities present. With the exception of BF_3 and C_2N_2, the above mentioned gases and CH_2F_2 have been generously provided by I.C.I. Ltd., who have also reported on the impurities. It is concluded that the major impurities are as listed in Table 1, where the percentages of impurities are given in some cases; in the others the quantities of contaminants are much less. The mass spectrometer studies, for example, show that CCl_3^+, the most abundant positive ion from CCl_4, where its appearance potential is 11·8 eV is negligible in CCl_3F at the latter voltage. Hence the CCl_3F did not contain appreciable amounts of CCl_4. CCl_3^+ is also found in CCl_3F but its appearance potential is then 13·8 eV. Studies of this kind show that impurities are negligible for the present series of breakdown strengths, made to the accuracy stated above. The most important impurities are I and Br in CIF_3 and

$CBrF_3$ respectively since negative ion formation in these elementary gases is very probable. Further, ClF_3 is liable to dissociate even on exposure to daylight and thus gives probably the most unreliable results of the present series, although the value quoted is not inconsistent with the other members of the series CF_3X.

Table 1. *Impurities in the test gases*

Gas	*Impurities*
CCl_3F	CF_2Cl_2 CCl_4
CCl_2F_2	CF_3Cl $CFCl_3$, CCl_4
$CClF_3$	CF_2Cl_2, CF_4
CF_4	C_2F_6 (1–2%) C_3F_8
$CBrF_3$	CHF_3 (2–3%) Br
ClF_3	CHF_3 (2–3%) I (0·5–1%)
SF_6	S_2F_{10} SF_4

Table 2. *Breakdown voltages measured in various gases between 2·0 cm diameter spheres, spaced 1 mm apart*

Gas	*Pressure (mm Hg)*	*Temperature (°C)*	*Relative breakdown strength c.f. nitrogen at the same pressure*	*Previous value of relative breakdown strength*	*Reference*
CCl_3F	630	16	$2·6_5$	3·0 4·5	(10) (12)
CCl_2F_2	815	20	$2·0_5$	2·4 2·6 2·1	(10) (12) (11)
	450	20	$1·8_7$		
	250	20	$1·8_0$		
$CClF_3$	815	20	$1·2_2$	—	—
CF_4	812	18	$0·9_2$	1·0 1·1	(11) (13)
$CBrF_3$	450	17	$1·4_1$	—	—
ClF_3	450	17	$2·0_5$	—	—
BF_3	827	20	1·0	—	—
Cl_2	817	14	$1·5_1$	$1·5_5$	(10)
SF_6	813	17	$1·8_5$	1·9 2·5	(11) (12)
C_2N_2	250	20	2·2	about 6	(15)
H_2	822	19	$0·5_4$	0·6	(18)

EXPERIMENTAL RESULTS

A selection of the data is given in Table 2, together with some data of other authors, taken in comparable conditions. The relative breakdown strengths listed in Table 2 are the ratios of the breakdown voltage for the gas considered to the breakdown voltage in nitrogen at the same pressure and temperature and at 1 mm gap spacing. It will be noted that the relative dielectric strengths of CF_3I, CF_3Br, CF_3Cl and CF_4 decrease in that order; this is also the order of decreasing negative ion yield, as measured in a mass spectrometer at pressures $\simeq 10^{-5}$ mm of mercury. Kowalenko[14] found the same trend of dielectric strength with HI, HBr and HCl.

C_2N_2 was studied because Natterer in his very early work[15] on sparking voltages in gases found it, with his extremely inaccurate methods, to have some six times the dielectric strength of N_2. It was thus considered worthwhile, in view of the internal consistency of many of Natterer's results, to check the C_2N_2 result, which (Table 2) we do not confirm. C_2N_2 has been the subject of detailed study in our mass spectrometers and the negative ion yield is qualitatively consistent with the measured breakdown strength given in Table 2.

Since Charlton and Cooper[10] had found a very low breakdown strength for CH_2F_2, this was also checked and found to give (in the units of Table 2) at 823 mm of mercury (18° C), 3 mm spacing between 2 cm spheres, a relative dielectric strength of 0·73, to be compared with Charlton and Cooper's result of 0·69. Again, mass spectrometer studies show only a very small yield of negative ions. These results on CH_2F_2 confirm those with CF_4 in that organic molecules containing only F (see, however, the discussion of SF_6 below) as the halogen component tend to give relatively low breakdown strengths, even though the production of free F by dissociation should be appreciable. It has been shown, for example, that the reaction

$$CF_4 + e \rightarrow CF_3^+ + F + 2e$$

is by far the most probable ionization process in CF_4. This again suggests that the formation of free neutral halogen atoms by dissociation is not of great importance in determining their dielectric strengths, although it has sometimes been implied. The results with SF_6 are, here, of great interest in that the negative ion yield in the mass spectrometer is much greater than with CF_4 and, further, consists mainly of SF_6^-, whereas F^- predominates in the CF_4 dissociation pattern. These differences cannot yet be fully explained. The high dielectric strength ($\times 4·5$ that of N_2) of SeF_6 reported by Hochberg and Sandberg[12] has not been confirmed, since

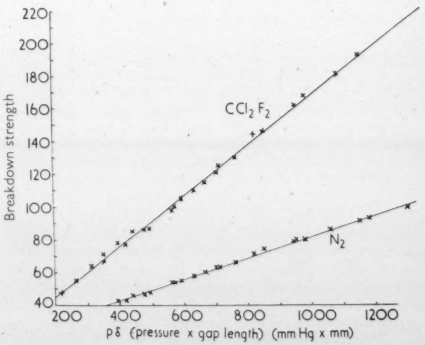


Fig. 1. Breakdown voltage data for temperature, 20° C. Breakdown strengths are in arbitrary units

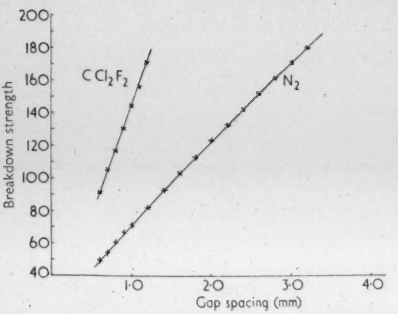


Fig. 2. Breakdown voltage for $p = 815$ mm of mercury; temperature, 20° C. Breakdown strengths are in arbitrary units

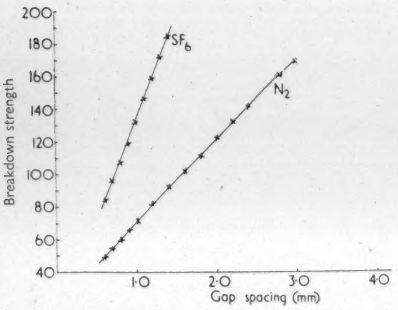


Fig. 3. Breakdown voltage data for $p = 813$ mm of mercury; temperature, 17° C. Breakdown strengths are in arbitrary units

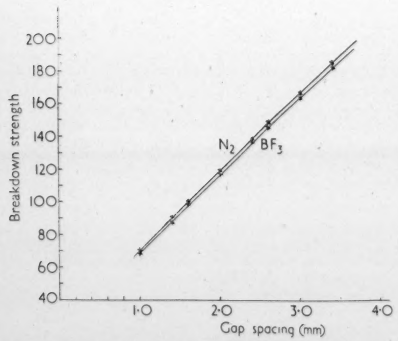


Fig. 4. Breakdown voltage data for $p = 827$ mm of mercury; temperature, 20° C. Breakdown strengths are in arbitrary units

Banks and Rudge[17] have recently found the dielectric strength figure of $\times 2{\cdot}9$ that of N_2.

It is next of interest to speculate briefly on the possibility of finding gases with higher breakdown strengths than CCl_2F_2 and SF_6 which are the most commonly used gases of this type in technical practice. The main requirements are a high vapour pressure at room temperature and reasonable stability. Molecules containing I are likely to provide, other things being equal, the highest breakdown strengths but their chemical stability tends to be less than that of similar molecules containing F or Cl. It thus appears unlikely that iodine compounds will be found more suitable than CCl_2F_2 and SF_6. Investigation of bromine compounds might be profitable.

Finally, the obvious limitations of the data of Table 2 should be briefly indicated. They refer only to the pressures and gap length stated, but Paschen's law applies, as shown in Fig. 1, for CCl_2F_2 and nitrogen, so that relative dielectric strengths may be found for pressures and gap lengths other than those used in the tests, within limits.

The variation of breakdown strength with spark gap length is shown in Figs. 2, 3 and 4 for CCl_2F_2, SF_6 and BF_3.

CONCLUSIONS

The measurements of breakdown strengths of various gases for a sphere gap with spacings of a few mm confirm earlier data in general. New measurements, made to provide preliminary information for a detailed study of electrical breakdown processes in polyatomic molecular gases, show that SF_6 and CCl_2F_2 still appear to be the most convenient and useful for insulation purposes.

REFERENCES

(1) MOHR, E. I., and WEISSLER, G. L. *Phys. Rev.*, **72**, p. 294 (1947).
(2) PENNING, F. M. *Ned. Tijd voor Natuurkunde*, **5**, p. 33 (1938).
(3) LLEWELLYN JONES, F., and PARKER, A. B. *Proc. Roy. Soc. A*, **213**, p. 185 (1952).
(4) DUTTON, J., HAYDON, S. C., and LLEWELLYN JONES, F. *Proc. Roy. Soc. A*, **213**, p. 203 (1952).
(5) LOEB, L. B. *Brit. J. Appl. Phys.*, **3**, p. 341 (1952).
(6) WARREN, J. W., HOPWOOD, W., and CRAGGS, J. D. *Proc. Phys. Soc.* [*London*] *B*, **63**, p. 180 (1950).
(7) WARREN, J. W., MARRIOTT, J., and CRAGGS, J. D. *Nature* [*London*], **171**, p. 514 (1953).
(8) BRUCE, F. M. *J. Instn Elect. Engrs*, **94**, II, p. 138 (1947).
(9) British Standard No. 358 (1939).
(10) CHARLTON, E. E., and COOPER, F. S. *Gen. Elect. Rev.*, **40**, p. 438 (1937).
(11) CAMILLI, G., and CHAPMAN, J. J. *Trans. Amer. Inst. Elect. Engrs*, **66**, p. 1463 (1947).
(12) HOCHBERG, B., and SANDBERG, E. *J. Tech. Phys. USSR*, **12**, p. 65 (1942).
(13) HOCHBERG, B. *Elektrichestvo*, No. 3, p. 15 (1947).
(14) KOWALENKO, G. M. *J. Tech. Phys. USSR*, **3**, p. 455 (1940).
(15) NATTERER, K. *Ann. Phys.* [*Leipzig*], **38**, p. 663 (1889).
(16) CRAGGS, J. D., MACDOWELL, C. A., and WARREN, J. W. *Trans Faraday Soc.*, **48**, p. 1093 (1952).
(17) BANKS, A. A., and RUDGE, A. J. *Nature* [*London*], **171**, p. 390 (1953).
(18) GEBALLE, R., and LINN, F. S. *J. Appl. Phys.*, **21**, p. 592 (1950).

Control systems with quasi-critical damping

By A. T. Fuller, M.A., Waymouth Gauges and Instruments Ltd., Godalming, Surrey

[*Paper first received* 11 *August, and in final form* 30 *December*, 1953]

The concept of critical damping is extended to linear systems of order greater than two, and is then termed "quasi-critical damping." When applied to a simple class of feedback systems, this condition is shown to have various simple interpretations in terms of (i) the poles of the closed-loop transfer function, (ii) the coefficients of the characteristic equation, or (iii) the open-loop lags. For practical systems, one of the open-loop lags must predominate over the others, and a very straightforward design rule may then be applied. Quasi-critical damping thus constitutes an appropriate design criterion of stability, particularly in the early theoretical stages of system synthesis.

1. Introduction

The stability of control systems is often assessed in terms of such concepts as "gain-margin," "phase-margin" or "amplitude-response at resonant frequency."[1,2] These have the merit of being easy to interpret on the Bode plot of amplitude against log-frequency or on the Nyquist diagram; and, from the designer's viewpoint, are especially convenient when a system is already partially synthesized and requires only minor modifications. When, however, a system is in the early theoretical stages of design, some non-graphical criterion of stability is desirable. The above concepts are then no longer suitable since they lead to cumbersome expressions when interpreted algebraically.

On the other hand, very simple algebraic criteria may be obtained by considering the case of zero damping (i.e. the condition when oscillations are just sustained), as has been shown by Prinz.[3] This treatment gives a useful guide to the designer, but actual zero damping is not, of course, a practicable condition for most control systems.

Another condition of stability is that of critical damping, and it is the purpose of the present paper to show that this, when suitably generalized for high-order systems, has the algebraic simplicity of the zero damping condition. Since it retains something of the direct applicability of the above graphical concepts, "quasi-critical damping" therefore forms a useful design criterion of stability.

2. Definition of Quasi-Critical Damping

Critical damping is a familiar concept when used to describe simple second-order systems, such as galvanometer movements. An intuitively acceptable definition might be along the following lines: definition *A*: "the response to a step-change of input is such that the output just fails to overshoot its final equilibrium value."

However, this definition turns out to be difficult to interpret algebraically, and may be ambiguous when applied to high-order systems. There is an alternative property of critically damped second-order systems which is more suitable for generalizing, namely: when damping is slightly reduced below critical, the overshoot following a disturbance does not merely occur once, but *continually recurs*, as time increases to infinity. A definition deriving from this property is as follows: definition *B*: "the response to a step-change of input is such that the output eventually reaches a regime in which it just fails to overshoot its final equilibrium value, this regime extending to infinite time."

When definitions *A* or *B* refer to systems of order greater than two, the condition will be termed "quasi-critical damping."

As it is outside the scope of the present paper to discuss further the relative merits of the two definitions, results will be given for both. Definition *B* will be used initially to obtain algebraic conditions, and it will then be shown graphically that, for the simple systems considered here, *the same conditions also satisfy definition A*.

3. Description of System

The system to be considered consists of a number of simple lags and a lagless amplifier arranged in a single control loop, with full feedback from output to input. An equivalent electrical circuit is shown in Fig. 1. In mathematical terms the system may be described as one with an open-loop

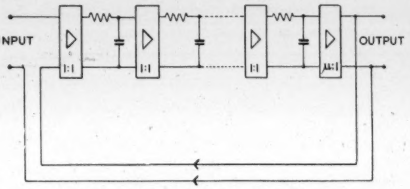


Fig. 1. Equivalent electrical circuit of the control system

transfer function having only real negative poles and no zeros, and with a feedback ratio of unity. These restrictions, which are made to simplify the calculations, imply a relatively elementary type of system, though one of widespread application. They do not exclude from the loop the possibility of (*a*) an integrator (which may be represented by a lag with infinite time-constant), or (*b*) a pure delay (which may be represented by an infinite number of infinitesimal lags).

It may be as well to mention that many of the subsequent results are also valid for more complicated systems, but it is not proposed to investigate their exact scope in the present paper.

4. Basic Condition for Quasi-Critical Damping

The open-loop transfer function $F(p)$ for the system defined in Section 3 is given by:

$$F(p) = \mu/[(1 + pT_1)(1 + pT_2) \ldots (1 + pT_n)] \qquad (1)$$

where μ is the amplifier gain; $T_1, T_2, \ldots T_n$ are the time-constants of the lags; and p is the Heaviside or Laplace operator representing differentiation.

The closed-loop transfer function $G(p)$ is given by:

$$G(p) = [1 + 1/F(p)]^{-1} \quad (2)$$

From equations (1) and (2):

$$G(p) = \mu/[\mu + (1 + pT_1)(1 + pT_2) \ldots (1 + pT_n)] \quad (3)$$

Let the poles of $G(p)$ be $P_1, P_2, \ldots P_n$. Then equation (3) may be written:

$$G(p) = \mu/[T_1T_2 \ldots T_n(p - P_1)(p - P_2) \ldots (p - P_n)] \quad (4)$$

When a step-change is applied to the input, the output $V(t)$ is given by:

$$V(t) = C_0 + C_1 \exp(P_1t) + C_2 \exp(P_2t) + \ldots + C_n \exp(P_nt) \quad (5)$$

where t is time, and the constants $C_0, C_1, \ldots C_n$ may be calculated from equation (4) by simple Laplace transform theory (see Chestnut and Mayer[2] for a lucid exposition of the method).

For the system to be stable, the exponential terms in equation (5) die out as time increases, and ultimately $V(t)$ approaches its equilibrium value C_0. The difference between $V(t)$ and C_0 at any instant will be called the error $E(t)$. Also, let the real and imaginary parts of $P_1, P_2, \ldots P_n$ be $-\alpha_1, -\alpha_2, \ldots -\alpha_n$ and $\beta_1, \beta_2, \ldots \beta_n$ respectively. Then equation (5) becomes:

$$E(t) = C_1 \exp[(-\alpha_1 + j\beta_1)t] + C_2 \exp[(-\alpha_2 + j\beta_2)t] + \ldots + C_n \exp[(-\alpha_n + j\beta_n)t] \quad (6)$$

Note that $\alpha_1, \alpha_2, \ldots \alpha_n$ must all be positive since the system is considered as being in a stable condition. Let them be suffixed in order of magnitude, i.e.

$$0 < \alpha_1 \leqslant \alpha_2 \leqslant \alpha_3 \ldots \leqslant \alpha_n \quad (7)$$

To find the condition for equation (6) to represent a quasi-critically damped response, its behaviour at large values of time must be examined, according to definition *B* in Section 2. It is convenient to distinguish between the following three cases:

(i) $\beta_1 = 0$ and $\alpha_1 < \alpha_2$

(ii) $\beta_1 \neq 0$

(iii) $\beta_1 = 0$ and $\alpha_1 = \alpha_2$

Case (i). Substituting $\beta_1 = 0$ in equation (6), and rearranging:

$$E(t) = \exp(-\alpha_1t)\{C_1 + C_2 \exp[(\alpha_1 - \alpha_2 + j\beta_2)t] + C_3 \exp[(\alpha_1 - \alpha_3 + j\beta_3)t] + \ldots + C_n \exp[(\alpha_1 - \alpha_n + j\beta_n)t]\} \quad (8)$$

Now from equation (7), $(\alpha_1 - \alpha_2) < 0, (\alpha_1 - \alpha_3) < 0, \ldots, (\alpha_1 - \alpha_n) < 0$.

Therefore, when $t \to \infty$, all the exponential terms in the braces $\{\}$ of equation (8) die out.

Therefore $E(t) \to C_1 \exp(-\alpha_1t)$

The error $E(t)$ thus eventually becomes a single-signed function, i.e. remains either continuously positive or continuously negative as time increases. Therefore, in this case, there is *ultimately* no overshoot.

Case (ii). In this case P_1 is complex, and one of the other poles, say P_2, must be its conjugate. Hence putting $\alpha_2 = \alpha_1$ and $\beta_2 = -\beta_1$ in equation (6) and rearranging:

$$E(t) = \exp(-\alpha_1t)\{C_1 \exp(+j\beta_1t) + C_2 \exp(-j\beta_1t) + C_3 \exp[(\alpha_1 - \alpha_3 + j\beta_3)t] + \ldots + C_n \exp[(\alpha_1 - \alpha_n + j\beta_n)t]\} \quad (9)$$

Assuming for the moment that $\alpha_3 \neq \alpha_1$, it can be seen by the method used in Case (i), that all except the first two exponential terms in the braces $\{\}$ in equation (9) die out as $t \to \infty$. Thus:

$$E(t) \to \exp(-\alpha_1t)D \cos(\beta_1t + \theta) \quad (10)$$

where D and θ are constants which may easily be calculated in terms of C_1 and C_2. As t increases, the cosine term in equation (10) becomes alternately positive and negative, and hence so does the error $E(t)$. Therefore, in this case, overshoot continually recurs.

[It is easily verified that this result remains valid when $\alpha_3 = \alpha_1$, and also when other α's equal α_1, as these cases merely correspond to introducing more cosine terms in equation (10)].

Case (iii). This case is intermediate between the others, in the sense that an infinitesimal change of α_1 or β_1 will change a system from Case (iii) to Case (i) or (ii) respectively. Case (iii) therefore represents the boundary between continual-overshoot and no-eventual-overshoot, i.e. according to definition *B* (in Section 2), Case (iii) represents the condition for quasi-critical damping.

The results of the three cases may be summarized by the following statement:

A system is quasi-critically damped, according to definition B, when two or more of its poles have the same least value of decay-rate, and one or more of these poles is real and negative.

5. INTERPRETATION IN TERMS OF CHARACTERISTIC EQUATION COEFFICIENTS

As shown in the last section, quasi-critical damping is determined by the nature of the poles of $G(p)$. These poles are the roots of the *characteristic equation*, formed by equating the denominator in equation (3) to zero. It is well known that if such an equation is more than second-order, the calculation of the roots is tedious, and if the order is more than four, a general literal solution is not available. We therefore attempt to express the condition for quasi-critical damping in terms of the coefficients of the polynomial in p obtained by multiplying out the characteristic equation. (This approach is analogous to finding conditions for zero damping by means of the Routh–Hurwitz criteria.[4,5])

The characteristic equation is $f(p) = 0$, where from equation (3):

$$f(p) = \mu + (1 + pT_1)(1 + pT_2) \ldots (1 + pT_n) \quad (11)$$

or, expressed as a polynomial in p:

$$f(p) = a_0 + a_1p + a_2p^2 + \ldots + a_np^n \quad (12)$$

where $a_0, a_1, \ldots a_n$ are constants which may be readily evaluated. If $f(p)$ is plotted against p (following Callender, Hartree and Porter[6]), with μ given the value zero, a graph is obtained similar to Fig. 2. This intersects the p-axis at the roots of $f(p) = 0$, which on putting $\mu = 0$ in equation (11)

are seen to be $-1/T_1$, $-1/T_2$, . . . $-1/T_n$. It is here supposed that $T_1, T_2, \ldots T_n$ are suffixed in order of magnitude, i.e.

$$T_1 \geqslant T_2 \geqslant T_3 \ldots \geqslant T_n \tag{13}$$

Now let μ be increased from zero to some small positive value. The effect is obviously to move the graph bodily in the direction of the vertical axis. Thus the two roots of smallest magnitude are no longer $-1/T_1$ and $-1/T_2$, but

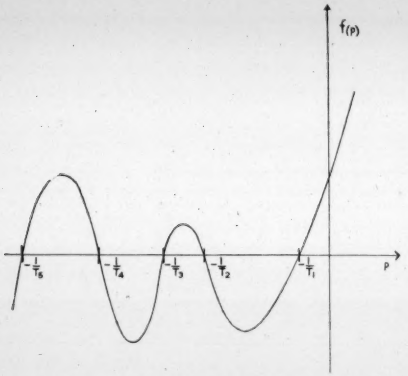


Fig. 2. Graph of $f(p)$ against p when $\mu = 0$

some values intermediate between the latter. As μ is further increased the two roots approach one another and become equal; and for still further increase of μ they become complex. According to Section 4, quasi-critical damping is reached when the two roots are equal, i.e. when the graph touches the p-axis tangentially. At the point of touching we have two conditions:

$$\left.\begin{aligned} f(p) &= 0 \\ \text{and} \quad f'(p) &= 0 \end{aligned}\right\} \tag{14}$$

or, from equations (12) and (14):

$$a_0 + a_1 p + a_2 p^2 + \ldots + a_n p^n = 0 \tag{15}$$

and

$$a_1 + 2a_2 p + \ldots + n a_n p^{n-1} = 0 \tag{16}$$

Elimination[7] of p from equations (15) and (16) gives

$$\begin{vmatrix} a_1 & 2a_2 & 3a_3 & \ldots & \ldots & 0 \\ a_0 & a_1 & a_2 & \ldots & \ldots & 0 \\ 0 & a_1 & 2a_2 & \ldots & \ldots & 0 \\ 0 & a_0 & a_1 & \ldots & \ldots & 0 \\ 0 & 0 & a_1 & \ldots & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & \ldots & \ldots & a_{n-1} & a_n \\ 0 & \ldots & \ldots & (n-1)\,a_{n-1} & & n a_n \end{vmatrix} = 0 \tag{17}$$

This is the condition which the coefficients $a_0, a_1, \ldots a_n$ must satisfy for the system to be quasi-critically damped. (Solving the determinant will give a polynomial in μ. One root of this will give the value of μ for quasi-critical damping—the other roots being of little interest since they correspond to values of μ when pairs of the larger magnitude roots of $f(p) = 0$ become equal.)

6. INTERPRETATION IN TERMS OF OPEN-LOOP LAGS

The expression for quasi-critical damping given by equation (17) is susceptible to a simple interpretation in terms of the lags distributed round the control loop. This is perhaps best shown by considering four separate systems, in order of increasing complexity.

System (a). This system has only two lags, T_1 and T_2, and provides a simple and easily verifiable example. From equations (11) and (12), the characteristic equation is

$$f(p) = \mu + (1 + pT_1)(1 + pT_2) \tag{18}$$

and is also

$$f(p) = a_0 + a_1 p + a_2 p^2 \tag{19}$$

Hence from equations (17) and (19), the condition for quasi-critical damping is

$$\begin{vmatrix} a_1 & 2a_2 & \cdot \\ a_0 & a_1 & a_2 \\ \cdot & a_1 & 2a_2 \end{vmatrix} = 0$$

or

$$a_1^2 = 4a_0 a_2 \tag{20}$$

Incidentally, equation (20) represents the well-known condition for equation (19) to have equal roots, so the determinant is verified for this particular case.

Substituting in equation (20) the values obtained by comparing coefficients in equations (18) and (19); writing μ_c as the value of μ for quasi-critical damping; and rearranging:

$$\mu_c = \tfrac{1}{4}[(T_1/T_2) - 2 + (T_2/T_1)] \tag{21}$$

System (b). Consider next a system with three lags: T_1, T_2 and T_3.

The condition for quasi-critical damping corresponding to equation (17) is

$$\begin{vmatrix} a_1 & 2a_2 & 3a_3 & \cdot & \cdot \\ a_0 & a_1 & a_2 & a_3 & \cdot \\ \cdot & a_1 & 2a_2 & 3a_3 & \cdot \\ \cdot & a_0 & a_1 & a_2 & a_3 \\ \cdot & \cdot & a_1 & 2a_2 & 3a_3 \end{vmatrix} = 0 \tag{22}$$

Expanding this determinant, and substituting for a_0, a_1, a_2, a_3, leads to a result which, for compactness, will be expressed in terms of *ratios* of the lags. Thus, writing:

$$M = T_1/(T_2 + T_3) \equiv \text{ratio of the } \textit{major lag} \text{ to the sum of the } \textit{minor lags}$$

$$N = T_2/T_3 \equiv \text{ratio of the minor lags}$$

and, further, writing:

$$K = N + 2 + (1/N)$$

we obtain from equation (22):

$$\mu_c = (27KM^2)^{-1}[2(K^2M^2 - KM - 3KM^2 + 1)^{3/2} - 2K^3M^3 + 3K^2M^2 + 9K^2M^3 + 3KM + 9KM^2 - 2] - 1 \tag{23}$$

Straightforward examination of equation (23) yields the

following results: if N is varied, keeping M constant, μ_c is maximum when $N = 1$, and decreases uniformly to equal minimum values at $N = 0$ and at $N = \infty$. This behaviour is illustrated in Fig. 3, which is a plot of μ_c against N, with M given the value 100.

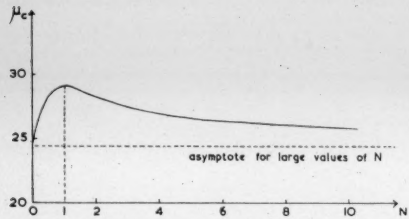


Fig. 3. Graph of quasi-critical gain against the ratio of the minor lags, for a 3-lag system, with $M = 100$

The minimum value $\mu_{c(min)}$ thus occurs when one lag becomes zero, i.e. when the system degenerates to the 2-lag case. In fact:

$$\mu_{c(min)} = \tfrac{1}{4}M[1 - (1/M)]^2 \tag{24}$$

which checks with equation (21).

The maximum value $\mu_{c(max)}$ occurs when the two minor lags become equal, and is given by:

$$\mu_{c(max)} = (8/27)M[1 - (1/2M)]^3 \tag{25}$$

It follows from equations (24) and (25) that for μ_c to be reasonably large ($\mu_c > 5$, say, for systems with a worthwhile ability to correct errors), M must be large (>17).

Therefore the ratio $\mu_{c(max)}/\mu_{c(min)}$ is approximately 32/27, which is near unity.

Hence μ_c is almost independent of N.

System (c). We have shown that the 3-lag case is most different from the 2-lag case when its two minor lags are equal (as might be expected), and that the difference is small. It seems reasonable, therefore, to suppose that the 4-lag case will differ most from the 3-lag and 2-lag cases when its three minor lags are equal (e.g. if one of the three minor lags is comparatively small, the system approximates to the 3-lag case, and if one is comparatively large, the other two become negligible, and the system approximates to the 2-lag case). This argument may be applied plausibly to higher order cases, provided the behaviour of μ_c shows no marked changes with increase of order.

Let us therefore examine a system of $(s + 1)$th order, having its s minor lags all equal and of value T_2, its major lag being T_1. In this case it is simpler to use equations (14) than to refer to the determinant directly. These give:

$$\mu + (1 + pT_1)(1 + pT_2)^s = 0 \tag{26}$$

and

$$(T_1 + sT_2 + psT_1T_2 + pT_1T_2)(1 + pT_2)^{s-1} = 0 \tag{27}$$

Solving equation (27) for p, and substituting in equation (26), gives:

$$\mu_c = M\left[\left(1 - \frac{1}{sM}\right)\Big/\left(1 + \frac{1}{s}\right)\right]^{s+1} \tag{28}$$

where M is, as before, the ratio of the major lag to the sum of the minor lags and equals $T_1/(sT_2)$. (A trivial result, $\mu = 0$, is also obtained from equations (26) and (27). This is explained at the end of Section 5.)

Note that equation (28) checks with equations (24) and (25). Examination of equation (28) shows that μ_c increases with s, if M is kept constant. The maximum value occurs when $s \to \infty$, i.e. when the minor lags become equivalent to a *pure delay* (or "distance-velocity lag"), and is given by:

$$\mu_c = M \exp[-1 - (1/M)] \tag{29}$$

If, in this equation, M is varied, the ratio μ_c/M becomes maximum when $M \to \infty$, i.e. when the major lag becomes a *pure integrator*, and is then given by:

$$(\mu_c/M)_{max} = \exp(-1) \simeq 0{\cdot}3679 \tag{30}$$

Further, the minimum value of (μ_c/M) occurs when $s = 1$ in equation (28); and since, as explained for system (*b*), $\mu_c > 5$ for practical systems, the ratio is given by

$$(\mu_c/M)_{min} = 0{\cdot}23 \tag{31}$$

Comparing equations (30) and (31), it will be seen that (μ_c/M) is almost a constant, independent of s.

System (d). Let us now consider a general system, in which the minor lags are not necessarily equal. An exact formula for μ_c is out of the question since, as previously shown, the solution of a high-order polynomial equation is involved. Nevertheless, a simple approximate formula, of sufficient accuracy, may be derived as follows:

The curve in Fig. 2 is given by equation (11) with μ given the value zero, i.e. by:

$$y \equiv [f(p)]_{(\mu=0)} = (1 + pT_1)(1 + pT_2) \ldots (1 + pT_n) \tag{32}$$

Following the argument in Section 5, μ_c is given by the magnitude of the algebraic minimum value of y which occurs between $p = -1/T_1$, and $p = -1/T_2$. For any one value of p in this range, say $(-q)$, let any two minor lags, say T_3 and T_4, be varied; the other lags T_1, T_2, T_5, $T_6 \ldots T_n$ being constant. Equation (32) then becomes:

$$y = H[(1 - qT_3)(1 - qT_4)] \tag{33}$$

where H is a constant, and where

$$1/T_2 > q > 1/T_1 \tag{34}$$

From equation (33)

$$y = H[1 - q(T_3 + T_4) + q^2T_3T_4] \tag{35}$$

Now let the sum of the minor lags be fixed, so that $(T_3 + T_4)$ is fixed, and the only variable in equation (35) is (T_3T_4). Also, from equations (34) and (13), $qT_3 < 1$ and $qT_4 < 1$; therefore the square bracket in equation (33) is positive, and hence so is the square bracket in equation (35). Therefore y has greatest magnitude when (T_3T_4) is maximized, which, for $(T_3 + T_4)$ constant, is easily shown to occur when $T_3 = T_4$. This holds for all the possible values of q. Hence in Fig. 2 the curve for $T_3 = T_4$, between $(-1/T_1, 0)$ and $(-1/T_2, 0)$, lies outside the curves for all other combinations of T_3 and T_4. Therefore μ_c is greatest when $T_3 = T_4$.

This argument may be applied with any two of the minor lags considered as variable, instead of T_3 and T_4. Successive equalization of such pairs ultimately gives $T_2 = T_3 = T_4 = \ldots = T_n$ as the condition for the optimum value of μ_c.

Similarly, μ_c is minimized when T_3 and T_4 are made as unequal as possible, i.e. when one of them is made zero. Successive applications of this process show that an absolute

minimum of μ_c is reached when the total minor lag is concentrated in one simple lag.

Thus for a general system, μ_c must lie between the maximum and minimum values for a system with equal minor lags, given by equations (30) and (31) respectively. From the latter equations:

$$\mu_c/M = 0{\cdot}3 \text{ to within } \pm 25\%$$

i.e. for all practicable systems, a formula of sufficient accuracy for design purposes is:

$$\mu_c \simeq (0{\cdot}3)T_1/(T_2 + T_3 + \ldots + T_n) \qquad (36)$$

In general, therefore, the major lag (T_1) must considerably predominate over the other lags, for the system to have a useful degree of sensitivity. It also follows that for this class of control systems, not more than one integrator is permissible in the loop.

An important special case occurs when the system contains one integrator, i.e. when $T_1 \to \infty$. The sensitivity is then expressed by the ratio $(\mu_c/T_1) \equiv R$, known in process-control as "floating rate," and is given by

$$R \simeq (0{\cdot}3)/(T_2 + T_3 + \ldots + T_n) \qquad (37)$$

7. GRAPHS OF QUASI-CRITICAL RESPONSES

In the previous section various conditions have been formulated for quasi-critical damping, this being defined in terms of "no-eventual-overshoot" (definition *B*). It remains to investigate whether the same conditions satisfy definition *A*, i.e. give responses that never overshoot at any value of time.

Fig. 4 shows graphs of the responses to step-input for four different systems, each of which satisfies definition *B*. They all have the same total minor lag τ, but different numbers of separate minor lags. Graph *A* is for a system with only one minor lag, *B* is for two equal minor lags, *C* is for three equal minor lags, and *D* is for an infinite number of equal minor lags (i.e. a pure delay). In each system the major lag is infinite (corresponding to an integrator), since computation is simplified thereby, and as already shown, the major lag must be predominant in practical systems.

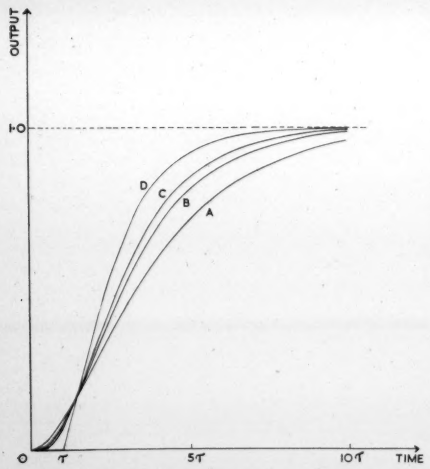


Fig. 4. Quasi-critically damped responses of four different systems to a unit-step input

Graphs *A*, *B*, *C* and *D* are for systems with one, two, three and infinity minor lags respectively; each having the sum of its minor lags equal to τ.

For each graph it will be seen that no overshoot takes place for values of time up to 10τ. It may readily be shown, for each graph, that by then the error is almost entirely due to the two equal poles with least decay-rate. Subsequently the fraction of the error due to the other poles decreases, and hence no overshoot takes place at any time.

We may infer that all the formulae given in Section 6 as satisfying definition *B* also satisfy definition *A*.

8. CORRELATION WITH THE LITERATURE

Following a study of open-loop response to step-input, Oldenbourg and Sartorious[8] assumed that many systems can be approximated by a pure delay combined with one simple lag. Their assumption is seen to be justified by the results of the present paper. They also gave implicitly equations (14) as representing the condition for "optimum response with aperiodic base-wave" (which is effectively quasi-critical damping according to definition *B*), and derived expressions equivalent to equations (29) and (30). Kun Li Chien, Hrones and Reswick[9] verified experimentally some results of Oldenbourg and Sartorious, and gave an empirical formula which checks with equation (36).

Prinz's[3] remark that one lag should predominate for a system to be critically damped is confirmed, and his further contention,[10] that many practical methods of stabilization amount fundamentally to achieving this predominance, is thought to be supported by the present paper.

9. CONCLUSIONS

An exact expression for quasi-critical damping of a simple class of control systems may be given in determinantal form, analogous to the relevant Routh–Hurwitz criterion. For practical systems, one open-loop lag must predominate, and a sufficiently accurate approximation is then: loop gain equals $(0{\cdot}3) \times$ ratio of major lag to sum of minor lags.

Since it is thus susceptible to exact calculation and can lead to simple results, quasi-critical damping constitutes an appropriate design criterion of stability.

It is hoped that these conclusions will stimulate attempts at more general and more rigorous treatments of the basic concept of quasi-critical damping.

ACKNOWLEDGEMENT

Thanks are due to the Directors of Waymouth Gauges and Instruments Ltd. for permission to publish this paper.

REFERENCES

(1) BROWN, G. S., and CAMPBELL, D. P. *Principles of Servomechanisms* (New York: John Wiley and Sons, Inc., 1948).

(2) CHESTNUT, H., and MAYER, R. W. *Servomechanisms and Regulating System Design*, Vol I (New York: John Wiley and Sons, Inc., 1951).

(3) PRINZ, D. G. *J. Sci. Instrum.*, **21**, p. 53 (1944).

(4) ROUTH, E. J. *Advanced Rigid Dynamics*, Vol. II (London: Macmillan and Co., Ltd., 1930).
(5) HURWITZ, A. *Mathematische Ann.*, **46**, p. 273 (1895).
(6) CALLENDER, A., HARTREE, D. R., and PORTER, A. *Phil. Trans. A*, **235**, p. 415 (1936).
(7) TURNBULL, H. W. *Theory of Equations* (London: Oliver and Boyd, Ltd., 1939).
(8) OLDENBOURG, R. C., and SARTORIOUS, H. *The Dynamics of Automatic Controls*. Translated by H. C. Mason (New York: American Society of Mechanical Engineers, 1948).
(9) KUN LI CHIEN, HRONES, J. A., and RESWICK, J. B. *Trans Amer. Soc. Mech. Engrs*, **74**, p. 175 (1952).
(10) PRINZ, D. G. Contribution to discussion of following paper:
WHITELY, A. L. *J. Instn Elect. Engrs*, **93**, II, p. 353 (1946).

Potential distribution of symmetrical cylindrical electron lenses

By G. D. ARCHARD, A.Inst.P., Associated Electrical Industries Ltd., Aldermaston, Berks.

[*Paper first received* 2 *September, and in final form* 23 *October*, 1953]

Expressions are developed for the distribution of axial potential and potential derivatives in the two-dimensional "unipotential" electron lens. Comparison is made with expressions obtained in simpler cases, and an example is illustrated.

The potential distribution and electron optical properties of rotationally symmetrical electron lenses have been widely treated in the literature. Less attention appears to have been paid to the two-dimensional ("cylindrical") electron lens, possibly on account of the relative lack of interest in the production of accurate line foci. This interest has recently been tending to increase, for example in connection with the correction of spherical aberration.(1) It therefore seems desirable to investigate fundamentally the properties of the two-dimensional analogues of existing rotationally symmetric electron lenses, in particular those of the electrostatic unipotential lens, which will now be considered. A first contribution to this end will be the determination of the potential distribution of a cylindrical lens represented by three collinear indefinitely long slits in three equally spaced infinitely large and infinitely thin planes, the lens being symmetrical about the centre plane and about the axis passing through the slits [Fig. 1(*b*)]. The outer planes, spaced by distances b from the centre plane at $x = 0$, may have potential V and the centre plane potential zero. The determination can be carried out more easily than in the case of rotational symmetry, for a potential varying in two dimensions lends itself, in simple cases such as that envisaged. to treatment by Schwarz' theorem. Applications of that theorem to potential problems are widespread; in electron optics, for example, it was applied by Glaser and Henneberg(2) to the case of a slotted diaphragm between two unbroken planes; the present arrangement is more general and would collapse into the former only in the case of very weak "pinhole" lenses, where $b \gg a, c$.

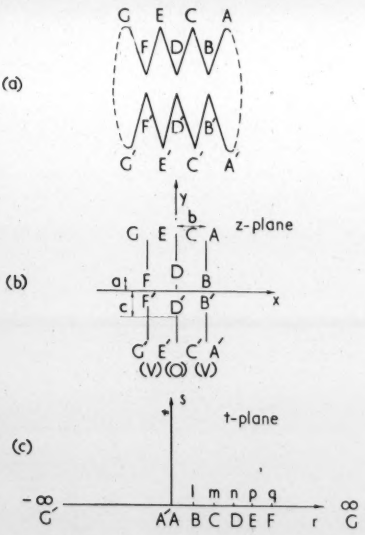


Fig. 1. Schwarz transformation for three slotted planes

TRANSFORMATION EQUATIONS

The principle of the Schwarz method is to treat the electrode array as the limiting case of a polygon [Fig. 1(*a*)] of which the angles γ and vertices (x, y) are as given in the table.

Co-ordinates of the polygon vertices

	x	y	γ	r
AA'	—	$\pm\infty$	2π	0
B	b	a	$-\pi$	1
C	—	∞	π	m
D	0	c	$-\pi$	n
E	—	∞	π	p
F	$-b$	a	$-\pi$	q
GG'	—	$\pm\infty$	2π	$\pm\infty$
F'	$-b$	$-a$	$-\pi$	$-q$
E'	—	$-\infty$	π	$-p$
D'	0	$-c$	$-\pi$	$-n$
C'	—	$-\infty$	π	$-m$
B'	b	$-a$	$-\pi$	-1

The inside of the polygon may be mapped on to the upper half of a t plane [Fig. 1(*c*)] by means of the transformation

$$z - z_0 = A \int \frac{(t^2 - 1)(t^2 - n^2)(t^2 - q^2)}{t^2(t^2 - m^2)(t^2 - p^2)} dt \qquad (1)$$

where $z = x + iy$ and $t = r + is$. The vertices of the

polygon transform into points $(r, 0)$ on the real axis of the t plane as indicated in the table.

Integration yields

$$(z - z_0) = A\left[t + \frac{n^2q^2}{m^2p^2t} + \frac{(m^2-1)(m^2-n^2)(m^2-q^2)}{2m^3(m^2-p^2)}\log\frac{t-m}{t+m} + \frac{(p^2-1)(p^2-n^2)(p^2-q^2)}{2p^3(p^2-m^2)}\log\frac{t-p}{t+p}\right] \quad (2)$$

Integration "half-round" C and E yields respectively

$$b = Ai\pi(m^2-1)(m^2-q^2)(m^2-n^2)/2m^3(m^2-p^2) \quad (3)$$

$$b = Ai\pi(p^2-1)(p^2-q^2)(p^2-n^2)/2p^3(p^2-m^2) \quad (4)$$

Insertion in equation (2) of the values of z for B, B', D, D', F, F' yields three identities and three equations relating a, b, c; thus

$$\frac{\pi a}{b} = \frac{2m^3(p^2-m^2)}{(m^2-1)(m^2-q^2)(m^2-n^2)}\left(1 + \frac{n^2q^2}{m^2p^2}\right) + \log\frac{(m+1)(p+1)}{(m-1)(p-1)} \quad (5)$$

$$\frac{\pi a}{b} = \frac{2m^3(p^2-m^2)}{(m^2-1)(m^2-q^2)(m^2-n^2)}\left(1 + \frac{n^2}{m^2p^2}\right)q + \log\frac{(q+m)(q+p)}{(q-m)(q-p)} \quad (6)$$

$$\frac{\pi c}{b} = \frac{2m^3(p^2-m^2)}{(m^2-1)(m^2-q^2)(m^2-n^2)}\left(1 + \frac{q^2}{m^2p^2}\right)n + \log\frac{(n+m)(p+n)}{(n-m)(p-n)} \quad (7)$$

{The order of the points $(1, m, n, p, q, \infty)$ must be remembered as it affects the log terms; e.g.

$$\log[(1-m)/(1+m)] = \log[(m-1)/(m+1)] + i\pi\}$$

These may be simplified by means of equations (3) and (4) which, together with (5) and (6) yield:

$$q = mp = n^2 \quad (8)$$

Equations (6), (7) and (8) now produce:

$$\frac{\pi a}{b} = \frac{2(m+p)(1+mp)}{(m^2-1)(p^2-1)} + \log\frac{(m+1)(p+1)}{(m-1)(p-1)} \quad (9)$$

$$\frac{\pi c}{b} = \frac{2(m+p)2\sqrt{mp}}{(m^2-1)(p^2-1)} + \log\frac{m+p+2\sqrt{mp}}{m+p-2\sqrt{mp}} \quad (10)$$

Equations (9) and (10) show that the problem is only concerned with the ratios a/b, c/b as one would expect. Thus, for particular values of a, b, c [Fig. 1(b)], the parameters m, n, p, q may be determined from equations (9) and (10) and substituted in equations (2) and (3). Elimination of A yields a transformation equation which puts points of the t half plane into one-to-one correspondence with points of the z plane, i.e.

$$t = t(z) = t(x + iy)$$

Now, if in some other complex plane $(\chi \equiv \phi + i\psi)$ a polygon be considered comprising simply the two straight lines $\phi = 0$ and $\phi = V$, points inside the polygon can be related to points of the t half plane by means of the transformation

$$\chi = i(V/\pi)\log[(t-l)/(t+l)] \quad (11)$$

and thence, by a reverse application of the former transformation, to points of the z plane, i.e.

$$\chi = \chi(z) = \chi(x + iy)$$

By a well known theorem (e.g. Zworykin and others[3]) the real part (ϕ) of χ obeys Laplace's equation in (x, y); if, therefore, it be arranged that the line $\phi = 0$ corresponds to the electrode(s) at zero potential [Fig. 1(b)] and the line $\phi = V$ to the electrode(s) at potential V, then the function

$$\phi = R(\chi) = \phi(x, y)$$

will give the potential distribution in the z plane.

Now if in equation (11) l be made equal to m, the line $\phi = 0$ will transform into $CDEFGG'F'E'D'C'$ and $\phi = V$ into $CBAA'B'C'$; but if l be made equal to p, the line $\phi = V$ will transform into $EFGG'F'E'$ and $\phi = 0$ into $EDCBAA'B'C'D'E'$. Neither case is desired. But the potential distributions obtained in the two cases (which may be called ϕ_m, ϕ_p respectively) both obey Laplace's equation in (x, y); hence so does their sum. Moreover, their sum gives the desired potentials on the electrodes ($\phi = 0$ on EDC and $C'D'E'$; $\phi = V$ on the remainder). Hence $\phi_m + \phi_p$ gives the entire potential distribution for the symmetrical lens. [Application requires care in dealing with $i\pi$ terms, as p has been assumed to be greater than m. The form of equation (12) allows for this fact.]

For ray tracing, only axial potentials and their derivatives are required. For axial potentials $r = 0$. The axial distribution is thus given by

$$\phi(x, 0) = \phi_m(x, 0) + \phi_m(-x, 0) \quad (12)$$

where $$\phi_m(x, 0) = (V/\pi)\tan^{-1}[2ms/(s^2 - m^2)] \quad (13)$$

and

$$x = -b + \frac{b}{\pi}\left\{\frac{2(m+p)(mp/s - s)}{(m^2-1)(p^2-1)} + \tan^{-1}\left[\frac{2ps}{(s^2-p^2)}\right] + \tan^{-1}\left[\frac{2ms}{(s^2-m^2)}\right]\right\} \quad (14)$$

The first and second derivatives of ϕ with respect to x are given by

$$\phi'(x, 0) = \phi'_m(x, 0) - \phi'_p(x, 0) \quad (15)$$

where $$\phi'_m(x, 0) = d\phi_m/ds \div dx/ds \quad (16)$$

$$\frac{dx}{ds} = -\frac{b}{\pi}\left[\frac{2(m+p)(mp/s^2+1)}{(m^2-1)(p^2-1)} + \frac{2p}{s^2+p^2} + \frac{2m}{s^2+m^2}\right] \quad (17)$$

$$\frac{d\phi_m}{ds} = -\frac{b}{\pi}\frac{2m}{s^2+m^2} \quad (18)$$

and $$\phi''(x, 0) = \phi''_m(x, 0) - \phi''_p(x, 0) \quad (19)$$

where $$\phi''_m(x, 0) = (d^2\phi_m/ds^2) \div (dx/ds)^2 - (d\phi_m/ds)(d^2x/ds^2) \div (dx/ds)^3 \quad (20)$$

$$\frac{d^2x}{ds^2} = \frac{2b}{\pi}\left[\frac{2(m+p)mp}{(m^2-1)(p^2-1)s^3} + \frac{2ps}{(s^2+p^2)^2} + \frac{2ms}{(s^2+m^2)^2}\right] \quad (21)$$

$$\frac{d^2\phi_m}{ds^2} = \frac{2b}{\pi}\frac{2ms}{(s^2+m^2)^2} \quad (22)$$

In use, the expressions are simpler than they appear, for m and p have numerical values and the right-hand sides of equations (18) and (22) are contained in equations (17) and (21).

To summarize: from the given geometry a, b, c, parameters m, p may be computed from equations (9) and (10). Then, by variation of s, equations (12), (13) and (15–22) give the axial potential and its first two derivatives corresponding to the axial co-ordinate given by equation (14).

SPECIAL CASES

For special relative values of a, b, c, equations (9) and (10) may be simplified. For example, if a/b is very small, both m and p must be large, and equation (9) may be rewritten

$$\pi a/b \simeq 2[(1/m) + (1/p)] \qquad (23)$$

If, at the same time, $c/b \gg a/b$, equation (10) may be written

$$\pi c/b \simeq \log [(m + p + 2\sqrt{mp})/(m + p - 2\sqrt{mp})] \qquad (24)$$

Unfortunately, the case of zero a/b and finite c/b (corresponding to that of Glaser and Henneberg[2]) cannot directly be reached, as both m and p would have to be infinite. If, however, p be made infinite and m remain finite, c/b will vanish and the electrode array will reduce to the already-known case of one unbroken plate and one slotted plate; equation (9) will reduce to

$$\frac{\pi a}{b} = \frac{2m}{m^2 - 1} + \log \frac{m + 1}{m - 1} \qquad (25)$$

(compare a particular case of the arrangement of Fry[4]).

If p be made equal to m, then c becomes infinite, that is, the central plate disappears and the expression for slot width becomes

$$\pi a/b = 4m(m^2 + 1)/(m^2 - 1)^2 + 2\log [(m + 1)/(m - 1)] \qquad (26)$$

This may easily be confirmed by direct application of Schwarz' theorem to the case of two slotted electrodes, although allowance must now be made for the fact that their potentials differ. It may be of value to write the expressions

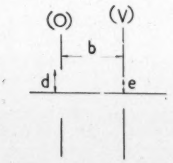


Fig. 2. Schwarz transformation for two slotted planes

for this case (Fig. 2) in full, without restriction to equality of slots, as follows:

$$\frac{\pi d}{b} = \frac{2mq(m^2 + 1)}{(m^2 - 1)(q^2 - m^2)} + \log \frac{q + m}{q - m} \qquad (27)$$

$$\frac{\pi e}{b} = \frac{2m(m^2 + q^2)}{(m^2 - 1)(q^2 - m^2)} + \log \frac{m + 1}{m - 1} \qquad (28)$$

$$x_{(y=0)} = \frac{b}{\pi}\left[\frac{2m^3(s - q^2/m^2 s)}{(m^2 - 1)(m^2 - q^2)} + \tan^{-1}\frac{2ms}{s^2 - m^2}\right] \qquad (29)$$

$$\phi_{(y=0)} = \frac{V}{\pi}\tan^{-1}\left[\frac{2ms}{(s^2 - m^2)}\right] \qquad (30)$$

It will be observed that both equations (27) and (28) collapse into equation (26) (apart from a constant factor) when q becomes equal to m^2 (compare equation (8) with m, n, p coincident).

EXAMPLE

In Fig. 3 are plotted ϕ, ϕ' and ϕ'' for the case $a = b = c = 1$, as given by the equations of the present work. For comparison, the potential distribution of the corresponding

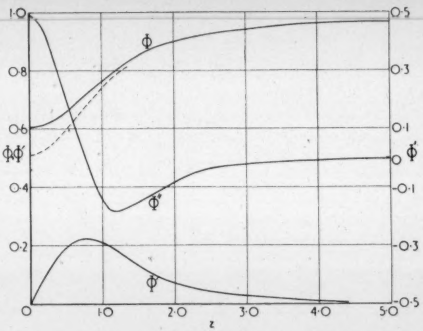


Fig. 3. Potential distribution of cylindrical lens

$a = b = c = 1$.
ϕ, ϕ'', symmetrical; ϕ', antisymmetrical.
– – – – – Regenstreif's values of ϕ for the corresponding rotationally symmetrical case.

rotationally symmetrical case as given by the theory of Regenstreif[5] is shown as a dotted line. The "dip" in the centre of the lens is greater for the latter case, as would be expected, but it is noticeable that Regenstreif's expression leads to a very slow flattening off of the potential at a distance from the lens.

ACKNOWLEDGEMENTS

I wish to thank Mr. M. E. Haine, Dr. G. Liebmann and Prof. D. Gabor for helpful comment, and Dr. T. E. Allibone for permission to publish this note.

REFERENCES

(1) SEELIGER, R. *Optik*, **8**, pp. 311–317 (1951).
(2) GLASER, A., and HENNEBERG, W. *Zeit. Techn. Phys.*, **16.8**, pp. 222–230 (1935).
(3) ZWORYKIN, V. K., MORTON, G. A., RAMBERG, E. G., HILLIER, J., and VANCE, A. W. *Electron Optics and the Electron Microscope*, p. 360 (London: Chapman and Hall, Ltd., 1945).
(4) FRY, T. C. *Amer. Math. Monthly*, **39**, p. 199 (1932).
(5) REGENSTREIF, E. *Ann. Radioélect.*, **6**, pp. 51–83 (1951).

Note added in proof. Since the acceptance of this paper, it has been found that similar work was carried out in 1952 by M. Laudet [*Cahiers de Phys.*, **41**, pp. 72–80 (1953)]. It is considered that there are sufficient differences in detail for the present paper to be published, particularly as it is intended to form the basis of a later paper on the focal properties of cylindrical lenses.

The emissivity of a groove

By L. F. Daws, M.Sc., The British Iron and Steel Research Association, London, W.1

[*Paper received* 1 *July*, 1953]

In order to check the performance of any temperature measuring device, which is dependent on radiation, a standard radiator should be available which possesses a thermal emissivity as close to unity as possible. One way the emissivity may be increased is to use a grooved target and in this paper an estimate is made of the improvement to be expected by using such a target. The variation of the effective emissivity with sighting angle is determined.

It is shown that a grooved target behaves as a Lambert surface when the angle of sighting is within a groove apex angle, but that this is no longer true for other sighting angles. The mean effective emissivity over all sighting angles is calculated using the relaxation method for a graphite target having a surface emissivity of 0·64 before grooving. After grooving it becomes 0·77, 0·82 and 0·82 for groove apex angles of 60°, 30° and 15° respectively. The corresponding emissivities when the angle of sighting is within the groove apex angle are 0·78, 0·88 and 0·94 respectively.

LIST OF SYMBOLS

e = target surface emissivity before grooving.

R_0 = initial radiation = $\sigma e T^4$ where T is the temperature of the groove surface.

$R(x)$ = rate of radiation from unit surface area of the groove at distance x from the apex.

$I(x)$ = integrated radiation function.

$R(\alpha)$ = $R(x)/R_0$ where $\alpha = x/a$ and a is the width of a groove face.

$$I(\alpha) = \int_0^{\alpha} R(\alpha)d\alpha.$$

R_e = total rate of radiation from unit length of groove.

R_1 = total rate if the surface emissivity before grooving is unity.

e_f = effective groove emissivity (a function of sighting angle ϕ).

e_{mf} = effective groove emissivity averaged over all sighting angles.

1. INTRODUCTION

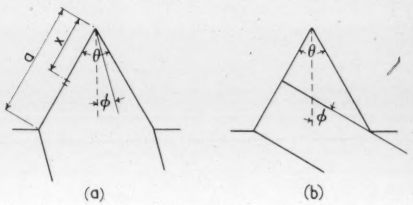
Any standard radiator must contain an aperture through which radiation emerges and consequently the effective emissivity of the radiator is reduced below unity. In order to compensate for this the graphite target of a radiator built at the British Iron and Steel Research Association Physics Laboratory[1] was ruled with grooves. The cavities so formed appreciably increase the effective surface emissivity of the target. In this paper the advantage to be obtained by grooving is theoretically estimated.

2. BEHAVIOUR OF A GROOVED SURFACE

It is assumed that a graphite surface behaves in the manner of a Lambert reflector to any incident radiant energy. That is to say, the radiation leaving an element of area within a small solid angle about any given direction is proportional to the projected area of the element in that direction. This radiation is equal to that from a black surface at the same temperature as the element times the element surface emissivity (e) plus (1–e) times the mean radiation per unit solid angle falling on to it. Thus the single quantity e describes the radiation properties of the surface.

When such a surface is ruled with regular grooves each element of the surface behaves as before, but a difficulty arises when an attempt is made to define the radiation properties of a target area which is large compared to the size of a groove. For the radiation per unit projected area of a flat surface is constant for all directions, which is not true of the ruled surface.

Suppose for simplicity that the groove is assumed infinitely long so that end effects are neglected, and let the width of each groove side be a. Then the mean radiation per unit solid angle per unit projected area of surface in a direction making an angle $\phi < \frac{1}{2}\theta$ [see Fig. 1(a)] with the plane



(a) (b)

Fig. 1. (a) Sighting angle (ϕ) within groove apex angle (θ)
(b) Sighting angle (ϕ) outside groove apex angle (θ)

bisecting the grooves (apex angle θ) over an area of surface is:

$$R_m = \frac{\left[\int_0^a \frac{R(x)}{\pi} \sin(\tfrac{1}{2}\theta - \phi)\,dx + \int_0^a \frac{R(x)}{\pi} \sin(\tfrac{1}{2}\theta + \phi)\,dx\right]}{2a \sin \tfrac{1}{2}\theta \cos\phi} \quad \phi < \tfrac{1}{2}\theta$$

where,

$R(x)$ is the rate of radiation from unit area of groove surface at distance x from the apex;

$dx \sin(\frac{1}{2}\theta - \phi)$ is the projected area of the element of unit length on one side of the groove;

$dx \sin(\frac{1}{2}\theta + \phi)$ is the projected area of the element of unit length on the other side; and

$2a \sin \frac{1}{2}\theta \cos\phi$ is the projected area of the equivalent flat surface across the mouth of the groove per unit length of groove.

$$R_m = \frac{\sin(\tfrac{1}{2}\theta - \phi) + \sin(\tfrac{1}{2}\theta + \phi)}{2a \sin \tfrac{1}{2}\theta \cos\phi} \int_0^a R(x)dx$$

$$= \frac{1}{a}\int_0^a R(x)dx \qquad (1)$$

which is independent of θ.

Thus for $\phi < \frac{1}{2}\theta$ the groove behaves as a Lambert surface. However, when $\phi > \frac{1}{2}\theta$ [see Fig. 1(*b*)]

$$R_m = \frac{1}{2a \sin \frac{1}{2}\theta \cos \phi} \int_b^a \sin (\tfrac{1}{2}\theta + \phi)\, R(x) dx \qquad (2)$$

where $b = [a \sin (\phi - \frac{1}{2}\theta)/\sin (\phi + \frac{1}{2}\theta)]$ and this is not independent of θ.

Thus the groove may only be considered a Lambert surface for radiation received and emitted within the groove apex angle. If radiation outside the angle is to be considered there is no such equivalent flat surface.

3. THE EFFECTIVE EMISSIVITY (e_f) AS A FUNCTION OF SIGHTING ANGLE

Because of the nature of the grooved surface its effective emissivity is difficult to define. The definition must depend on the use to which the target is put.

It may be used in such a way that the angle of sighting is within the groove apex angle when it is seen from equation (1) that

$$e_f = \frac{e}{aR_0} \int_0^a R(x) dx \qquad (3)$$

where, $R_0 = \sigma e T^4$

and T is the temperature of the groove surface.

Again, if the angle of sighting (ϕ) lies outside the apex angle, it has been shown in Section 2 that the target ceases to behave as a Lambert surface and the dependence of e_f on ϕ is found from equation (2) to be

$$e_f = \frac{e}{(a - b)R_0} \int_b^a R(x) dx \qquad (4)$$

where, $b = [a \sin (\phi - \frac{1}{2}\theta)/\sin (\phi + \frac{1}{2}\theta)]$

Where the grooved surface is to be used as the target of a black body furnace a mean emissivity over all sighting angles is required. For, although the sighting on the target will always be within a groove apex angle in this case, the emissivities over other sighting angles are required in order to determine the exchange of radiant energy between the walls of the furnace tube and the target. Buckley[2] has calculated this where the target behaves as a Lambert surface. Therefore, using Buckley's calculation and inserting a value for the target emissivity averaged over a hemisphere instead of the one valid only for instrument sightings, will enable a lower limit to be found for the blackness of the furnace.

4. MEAN EFFECTIVE EMISSIVITY (e_{mf}) OVER ALL SIGHTING ANGLES

This will be defined as the total radiation escaping from the groove divided by the radiation escaping if the groove surface emissivity is unity.

In Fig. 2 *ABC* represents a groove cross-section at right angles to the groove axis *AD*. With the co-ordinate system origin *O* as shown, and assuming that Lambert's cosine law is valid, the radiation from unit area of groove surface at (O,μ,ν), distance x from the apex, which passes through an element of area $d\xi d\eta$ in the groove base at (ξ,η,O) can be written in the form

$$\frac{R(x)}{\pi} \frac{\nu}{r} \frac{(\eta - \mu) \cos \frac{1}{2}\theta - \nu \sin \frac{1}{2}\theta}{r} \frac{d\xi d\eta}{r^2}$$

Thus the rate at which radiation passes out of the groove per unit area at x is

$$\frac{R(x)}{\pi} \int_{\eta = -a \sin \frac{1}{2}\theta}^{a \sin \frac{1}{2}\theta} d\eta \int_{-\infty}^{\infty} \frac{-\nu[(\eta - \mu) \cos \frac{1}{2}\theta - \nu \sin \frac{1}{2}\theta] d\xi}{r^4}$$

$$= \frac{R(x)}{2} \left[\frac{x - a \cos \theta}{(x^2 - 2ax \cos \theta + a^2)^{\frac{1}{2}}} + 1 \right] dx$$

where $r^2 = (\mu - \eta)^2 + \xi^2 + \nu^2$

Then, the total radiation escaping per unit length of groove is given by

$$R_e = \int_0^a \frac{R(x)}{2} \left[\frac{x - a \cos \theta}{(x^2 - 2ax \cos \theta + a^2)^{\frac{1}{2}}} + 1 \right] dx \qquad (5)$$

If the target emissivity is unity the total radiation is

$$R_1 = \int_0^a \frac{R_0}{2e} \left[\frac{x - a \cos \theta}{(x^2 - 2xa \cos \theta + a^2)^{\frac{1}{2}}} + 1 \right] dx$$

$$= \frac{aR_0}{e} \sin \tfrac{1}{2}\theta \qquad (6)$$

where R_0 is the initial radiation.

Thus the mean effective emissivity (e_{mf}) is given by

$$e_{mf} = \frac{R_e}{R_1} = \frac{e}{2a \sin \frac{1}{2}\theta} \int_0^a \frac{R(x)}{R_0} \left[\frac{x - a \cos \theta}{(x^2 - 2a\, x \cos \theta + a^2)^{\frac{1}{2}}} + 1 \right] dx \qquad (7)$$

i.e. in the normalized form

$$e_{mf} = \frac{e}{2 \sin \theta} \int_0^1 R(\alpha) \left[\frac{\alpha - \cos \theta}{(\alpha^2 - 2\alpha \cos \theta + 1)^{\frac{1}{2}}} + 1 \right] d\alpha \qquad (8)$$

where, $x = a\alpha$

$$R(x) = R_0 R(\alpha) \qquad (9)$$

It will be seen from equation (12) that this can be rewritten as

$$e_{mf} = \frac{e}{(1 - e)} \frac{[1 - eI(1)]}{\sin \frac{1}{2}\theta} \qquad (10)$$

It is easily seen from equations (3), (4) and (10) that in order to find the effective emissivity of a groove, whether the sighting angle is within or outside the apex angle, it is necessary to calculate the integrated radiation function, namely,

$$I(x) = \int_0^x R(\xi) d\xi$$

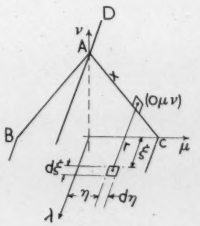


Fig. 2. Co-ordinate system used in Section 4

5. THE INTEGRAL EQUATION GIVING $I(x)$

The form of the relation which determines $R(x)$ can be written down by inspection. For the radiation coming from unit area of an element strip of width dx on one face of the groove (see Fig. 3) is equal to the initial radiation R_0, plus

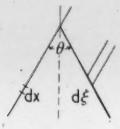


Fig. 3

reflected energy originating at all strips on the other face of the groove of width $d\xi$.

$$R(x) = R_0 + (1 - e)\int_0^a R(\xi) f(x,\xi) d\xi$$

where e is the surface emissivity of the groove and $f(x,\xi)$ depends on the solid angle subtended by the face of the groove containing strips $d\xi$, at unit area on the strip dx.

It is easily shown that

$$f(x,\xi) = \frac{x\xi \sin^2\theta}{2(x^2 - 2x\xi\cos\theta + \xi^2)^{\frac{1}{2}}}$$

so that

$$R(x) = R_0 + (1 - e)\int_0^a R(\xi)\frac{x\,\xi \sin^2\theta}{2(x^2 - 2x\xi\cos\theta + \xi^2)^{\frac{1}{2}}} d\xi \quad (11)$$

Integrating this equation with respect to x gives

$$I(x) = R_0 x + \frac{(1 - e)}{2}\int_0^a R(\xi)\left[\frac{x\cos\theta - \xi}{(x^2 - 2x\xi\cos\theta + \xi^2)^{\frac{1}{2}}} + 1\right] d\xi \quad (12)$$

$$\text{i.e.}\quad I(x) = R_0 x + \frac{1 - e}{2} I(a)\left[\frac{x\cos\theta - a}{(x^2 - 2ax\cos\theta + a^2)^{\frac{1}{2}}} + 1\right] + \frac{1 - e}{2}\int_0^a I(\xi)\frac{x^2\sin^2\theta}{(x^2 - 2x\xi\cos\theta + \xi^2)^{\frac{3}{2}}} d\xi \quad (13)$$

Writing this in its normalized form gives

$$I(\alpha) = + \frac{1 - e}{2} I(1)\left[\frac{\alpha\cos\theta - 1}{(\alpha^2 - 2\alpha\cos\theta + 1)^{\frac{1}{2}}} + 1\right] + \frac{1 - e}{2}\int_0^1 I(\beta)\frac{\alpha^2\sin^2\theta}{(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{3}{2}}} d\beta \quad (14)$$

where,

$$\left.\begin{aligned} I(\alpha) &= \frac{I(x)}{aR_0} \\ x &= a\alpha \\ \xi &= a\beta \end{aligned}\right\} \quad (15)$$

This then is the integral equation giving $I(x)$.

6. THE SOLUTION OF THE INTEGRAL EQUATION

Analytic solution. Equation (11) is of the Fredholm type with a symmetric nucleus[3] and therefore possesses a solution of the type

$$R(\alpha) = 1 + (1 - e)\int_0^1 f(\alpha,\beta)d\beta + (1 - e)^2\int_0^1 f(\alpha,\beta)d\beta\int_0^1 f(\alpha_1,\beta)d\alpha_1 + (1 - e)^3\int_0^1 f(\alpha,\beta)d\beta\int_0^1 f(\alpha_1,\beta)d\alpha_1\int_0^1 f(\alpha_1\beta_1)d\beta_1 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (16)$$

where, $f(\alpha,\beta) = \alpha\beta\sin^2\theta/2(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{1}{2}}$

It can be shown[4] that the above series converges for

$$|1 - e| < 1 \Big/ \left\{\int_0^1 d\alpha\int_0^1 [f(\alpha,\beta)]^2 d\beta\right\}^{\frac{1}{2}} \quad (17)$$

Now every term of the series after the second can be expressed in terms of the three standard elliptic integral forms which have been tabulated, but the amount of time consumed in doing this goes up exponentially with the number of terms required in the series. Even then a great deal of work is still necessary in order to tabulate $R(\alpha)$.

It was therefore decided to find particular solutions of the integral equation by an approximate means, using the Southwell relaxation technique.

7. THE RELAXATION METHOD OF SOLUTION

In order to solve equation (13) by a relaxation method, the integral on the right hand side of the equation is replaced by an approximation using the trapezoidal rule. That is the range of α is divided into n equal divisions. Then when α is equal to r/n equation (3.3) becomes

$$I\left(\frac{r}{n}\right) = \frac{r}{n} + \sum_{s=1}^{n-1} F(r,s) I\left(\frac{s}{n}\right) + \tfrac{1}{2}F(r,n)I(1) + f\left(\frac{r}{n}\right)I(1) \quad (18)$$

$$\left.\begin{aligned} \text{where}\quad F(\alpha,\beta) &= (1 - e)\frac{\alpha^2\sin^2\theta}{2(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{3}{2}}} \\ f(\alpha) &= \frac{(1 - e)}{2}\left[\frac{\alpha\cos\theta - 1}{(\alpha^2 - 2\alpha\cos\theta + 1)^{\frac{1}{2}}} + 1\right] \end{aligned}\right\} \quad (19)$$

In this way n linear algebraic relations may be formulated between the n unknowns $I(1/n)$ to $I(1)$. The solutions of this system of equations by the relaxation method is now quite straightforward. In fact the range of α was divided into ten equal parts and the resulting equations have been solved for $e = 0{\cdot}64$, and $\theta = 15°$, $30°$ and $60°$. The solutions with differences up to the third decimal place are given in Table 1.

It can be seen that the solutions are not particularly smooth near the origin. This is due to the presence of a pole of the integrand of equation (13) at $\alpha = 0$. The size of the region over which this effect is noticeable can be reduced by using a smaller interval. Thus for the case of $\theta = 30°$, the interval

Table 1. *Distribution of $I(\alpha)$ with α, for $e = 0{\cdot}64$ and $\theta = 15°$, 30° and 60°*

	$\theta = 15°$				$\theta = 30°$				$\theta = 60°$			
α	$I(\alpha)$	$\Delta I(\alpha)$	$\Delta^2 I(\alpha)$	$\Delta^3 I(\alpha)$	$I(\alpha)$	$\Delta I(\alpha)$	$\Delta^2 I(\alpha)$	$\Delta^3 I(\alpha)$	$I(\alpha)$	$\Delta I(\alpha)$	$\Delta^2 I(\alpha)$	$\Delta^3 I(\alpha)$
0	0	0·153	0·006	−0·012	0	0·147	0·008	−0·018	0	0·134	−0·001	−0·002
0·1	0·153	0·159	−0·006	0·007	0·147	0·155	0·010	−0·009	0·134	0·133	−0·003	−0·001
0·2	0·312	0·153	0·001	0·004	0·302	0·145	−0·001	−0·002	0·267	0·130	−0·004	0·001
0·3	0·465	0·154	−0·003	0·002	0·447	0·144	−0·003	−0·000	0·397	0·126	−0·003	0·001
0·4	0·619	0·151	−0·001	−0·003	0·591	0·141	−0·003	−0·001	0·523	0·123	−0·002	−0·001
0·5	0·770	0·150	−0·004	0	0·732	0·138	−0·004	−0·001	0·646	0·121	−0·003	0
0·6	0·920	0·146	−0·004	−0·003	0·870	0·134	−0·005	0	0·767	0·118	−0·003	0·002
0·7	1·066	0·142	−0·007	−0·003	1·004	0·129	−0·005	0	0·885	0·115	−0·001	−0·002
0·8	1·208	0·135	−0·010		1·133	0·124	−0·005		1·000	0·114	−0·003	
0·9	1·343	0·125			1·257	0·119			1·114	0·111		
1·0	1·468				1·376				1·225			

Table 2. *Recalculation of $I(\alpha)$ in the region $\alpha = 0$, for $\theta = 30°$, and interval 1/40*

α	1/40	2/40	3/40	4/40	5/40	6/40	7/40	8/40	9/40	10/40	11/40	12/40
$I(\alpha)$	0·037	0·076	0·111	0·151	0·189	0·226	0·263	0·300	0·337	0·373	0·410	0·446
$\Delta I(\alpha)$	0·039	0·035	0·040	0·038	0·037	0·037	0·037	0·037	0·036	0·037	0·036	
$\Delta^2 I(\alpha)$	−0·004	0·005	−0·002	−0·001	0	0	0	−0·001	0·001	−0·001		
$\Delta^3 I(\alpha)$	0·009	−0·007	0·001	0·001	0	0	−0·001	0·002	−0·002			

was reduced to a quarter and the solution obtained is given in Table 2.

8. THE VALUE OF THE RADIATION FUNCTION AT THE ORIGIN

There is no absolute check on the accuracy of the solutions given in Table 1 near the origin. However, it is possible to determine the value of the radiation function at the origin from equation (11), so that if it can be assumed that $I(\alpha)$ can be expressed in terms of a Taylor's series for small α

$$I(\alpha) = \alpha R(0) \tag{20}$$

Now expressing equation (11) in its normalized form

$$R(\alpha) = 1 + \frac{1-e}{2}\int_0^1 R(\beta)\frac{\alpha\beta\sin^2\theta}{(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{1}{2}}}d\beta \tag{21}$$

so that

$$R(0) = 1 + (1-e)\lim_{\alpha\to 0}\int_0^1 R(\beta)\frac{\alpha\beta\sin^2\theta}{(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{1}{2}}}d\beta \tag{22}$$

The integrand in equation (22) is zero for all values of β except $\beta = 0$. Therefore a small fixed quantity ϵ can be chosen such that

$$R(0) = 1 + (1-e)\lim_{\alpha\to 0}\int_0^{\epsilon} R(\beta)\frac{\alpha\beta\sin^2\theta}{2(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{1}{2}}}d\beta$$

In the remaining integral the integrand is positive over the range of integration. Therefore, using the mean value theorem for integrals

$$R(0) = 1 + (1-e)\lim_{\alpha\to 0} R(\xi)\int_0^{\epsilon}\frac{\alpha\beta\sin^2\theta}{2(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{1}{2}}}d\beta$$

i.e. $$R(0) = 1 + (1-e)\lim_{\alpha\to 0} R(\delta)\left[\frac{\beta\cos\theta - \alpha}{(\alpha^2 - 2\alpha\beta\cos\theta + \beta^2)^{\frac{1}{2}}}\right]_0^{\epsilon}$$

$$= 1 + \frac{1-e}{2}\lim_{\alpha\to 0} R(\delta)\left[\frac{\epsilon\cos\theta - \alpha}{(\alpha^2 - 2\alpha\epsilon\cos\theta + \epsilon^2)^{\frac{1}{2}}} + 1\right]$$

$$= 1 + \frac{1-e}{2}R(\delta)(1 + \cos\theta)\,0 \leqslant \delta \leqslant \epsilon \tag{23}$$

Since $R(\alpha)$ is a continuous function of α, and ϵ can be made as small as is necessary, equation (23) gives

$$R(0) = 1/[1 - (1-e)(1 + \cos\theta)/2] \tag{24}$$

so that from equation (20)

$$I(\alpha) = \alpha/[1 - (1-e)(1 + \cos\theta)/2] \tag{25}$$

for small α.

In Table 5 the values of $I(0{\cdot}1)$ obtained using this formula are compared with those obtained by relaxation. It can be seen that the differences are small. The solution for $I(0{\cdot}1)$ given in Table 2 agrees with the one given by equation (25). Since the value of $I(1)$ remained unaltered when the interval was reduced to a quarter, the solutions of Table 1 are quite accurate enough for determining effective emissivity.

9. THE SOLUTION FOR R(α)

The solution for $R(\alpha)$ has been obtained by using the relaxation method. To do this equation (12) was used in the form

$$\int_{\alpha_1}^{\alpha_2} R(\beta)d\beta = \alpha_2 - \alpha_1 + \frac{1-e}{2}\int_0^1 R(\beta)\left[\frac{\alpha_2\cos\theta - \beta}{(\alpha_2^2 - 2\alpha_2\beta\cos\theta + \beta^2)^{\frac{1}{2}}} - \frac{\alpha_1\cos\theta - \beta}{(\alpha_1^2 - 2\alpha_1\beta\cos\theta + \beta^2)^{\frac{1}{2}}}\right]d\beta \tag{26}$$

By dividing the range of β into n equal parts as before and replacing the integrals using the trapezoidal rule, n linear relations can be set up between the unknowns $R(1/n)$ to $R(1)$. $R(0)$ is known from equation (24).

The solution for $e = 0{\cdot}64$ and $\theta = 15°$, 30° and 60° are

Table 3. *Distribution of $\phi(\alpha)$ with α for $e = 0{\cdot}64$, and groove angles of 15°, 30° and 60°*

	$\theta = 15°$		$\theta = 30°$		$\theta = 60°$	
α	$\phi(\alpha)$	$\Delta\phi(\alpha)$	$\phi(\alpha)$	$\Delta\phi(\alpha)$	$\phi(\alpha)$	$\Delta\phi(\alpha)$
0	1·55	0	1·505	−0·01	1·37	−0·04
0·1	1·55	−0·02	1·50	−0·01	1·33	−0·01
0·2	1·53	0·01	1·49	−0·06	1·32	−0·05
0·3	1·54	0·01	1·43	0	1·27	−0·02
0·4	1·55	−0·07	1·43	−0·04	1·25	−0·03
0·5	1·48	0	1·39	−0·03	1·22	−0·03
0·6	1·48	−0·04	1·36	−0·05	1·19	−0·02
0·7	1·44	−0·06	1·31	−0·04	1·17	−0·02
0·8	1·38	−0·08	1·27	−0·05	1·15	−0·03
0·9	1·30	−0·09	1·22	−0·05	1·12	−0·01
1·0	1·21		1·17		1·11	

Table 4. *Effective target emissivity (e_{mf}) as a function of groove angle (θ)*

θ	$I(1)$ [*from $I(\alpha)$ distributions*]	*Check on* $I(1)$ [*from $\phi(\alpha)$ distributions*]	e_{mf}
15°	1·468	1·463	0·824
30°	1·376	1·373	0·819
60°	1·225	1·225	0·768

given in Table 5 for an interval of 1/10. It can be seen that they suffer from the same trouble as the $I(\alpha)$ solutions near the origin.

Table 5. *Comparison of $I(0{\cdot}1)$ obtained by relaxation with those obtained from the first item of the Taylor's series*

θ (*degrees*)	$I(0{\cdot}1)$ [*from equation* (25)]	$I(0{\cdot}1)$ (*from Table* 1)
60	0·155	0·153
30	0·151	0·148
15	0·137	0·134

10. RESULTS

It has been shown in this paper that a grooved target behaves as a Lambert surface when the sighting angle is within the groove apex angle, but that this is no longer true for sighting angles outside the apex angle.

However, for all sighting angles the effective emissivity (e_f) can be expressed in terms of the integrated radiation function, i.e.

$$I(\alpha) = \int_0^{\alpha} R(\alpha)d\alpha \tag{27}$$

where, $R(\alpha)$ is the rate of normal radiation from unit area of the grooved surface at distance x from the groove apex divided by the initial radiation R_0, and $\alpha = x/a$ where a is the width of a groove face. Thus when the sighting angle is within the groove apex angle θ

$$e_f = eI(1) \tag{28}$$

while when this is not so

$$e_f = e[I(1) - I(\alpha)]/(1 - \alpha)$$

where

$$\alpha = \sin(\phi - \tfrac{1}{2}\theta)/\sin(\phi + \tfrac{1}{2}\theta) \tag{29}$$

The average effective emissivity over all sighting angles (e_{mf}) (averaged over a hemisphere) which may be used to obtain a lower limit for the blackness of a black body furnace, is given by

$$e_{mf} = \frac{e[1 - eI(1)]}{(1 - e)\sin\tfrac{1}{2}\theta} \tag{30}$$

The distribution of $I(\alpha)$ and $R(\alpha)$ with α have been calculated using the relaxation method for groove angles of 15°, 30° and 60° and for a target emissivity before grooving of 0·64. They are given in Tables 1 and 3.

The variation of effective emissivity with sighting angle [calculated from equations (28) and (29)] is shown graphically in Fig. 4, while the mean effective emissivities calculated from equation (30) are given in Table 4.

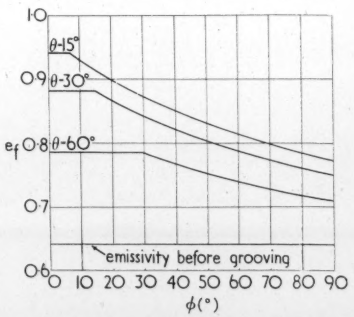


Fig. 4. Effective groove emissivity (e_f) as a function of sighting angle (ϕ) for apex angles (θ) = 15°, 30° and 60°

11. CONCLUSIONS

A considerable increase in the effective emissivity of a target can be obtained by grooving.

If the target behaves as a Lambert surface before grooving,

it will continue to do so after grooving for angles of sighting within a groove apex angle. For such sighting angles there is no limit to the degree of blackness theoretically, the effective emissivity tending to unity as the groove angle goes to zero. For groove angles of 60°, 30° and 15° the effective emissivities within a groove apex angle are respectively 0·785, 0·881 and 0·940, where the target surface emissivity before grooving was 0·64.

When the angle of sighting is outside a groove apex angle the target ceases to behave as a Lambert surface. The variation of effective emissivity with sighting angle is given in Fig. 4. It can be seen that e_f falls off quite slowly with ϕ and considerable advantage is to be obtained even with an angle of sighting close to 90°.

REFERENCES

(1) BURTON, E. J., and MAYORCAS, R. *J. Iron Steel Inst.*, **168**, p. 151 (1951).
(2) BUCKLEY, H. *Phil. Mag.*, **17**, p. 576 (1934).
(3) WHITTAKER, E. T., and WATSON, G. N. *A Course of Modern Analysis* (London: Cambridge University Press, 1946).
(4) RIEMANN-WEBERS. *Differentialgleichungen der physik*, p. 428–429 (Braunschweig: Vieweg, 1925).

The rolling resistance of a wheel with a solid rubber tyre

By I. EVANS, M.Sc., F.Inst.P., Fighting Vehicles Research and Development Establishment, Chertsey, Surrey

[*Paper received* 2 *September*, 1953]

The rolling resistance of a wheel with a solid rubber tyre, running on a rigid plane surface, is calculated on the assumption that it is due to hysteresis losses in cyclic deformation of the rubber. The results and certain other deductions from the theory are in reasonably good agreement with experimental observations.

In a paper on the mechanism of rolling friction Tabor[1] suggests that rolling friction in the elastic range is primarily due to elastic hysteresis losses. Independently of Tabor, though subsequent to his work on the subject, the author had used the same postulate to calculate the rolling resistance of a wheel with a solid rubber tyre.

THEORY

The problem is regarded as one in plane strain, the wheel being considered as a long composite cylinder consisting of a rigid core of radius a and a perfectly elastic deformable tyre of external radius b and width s. When the wheel is loaded by a weight W and itself bears on a rigid horizontal plane there is a zone of contact in which the pressure will increase from zero at the edges to a maximum along the centre line. Ideally, we should like to know the form of this pressure distribution and the distortion of the tyre produced by it, but there appears to be no tractable analytic solution of this problem. Accordingly the problem of the work done in deforming the tyre is tackled indirectly by making some rather sweeping assumptions, the merit of which must be judged from a comparison of the answer they give with experiment.

In the first place we assume that if p_0 is the average contact pressure and u the maximum vertical compression of the tyre, the work done in compressing unit area of the tyre through a distance u as the wheel turns is p_0u. If, owing to the resilience of the tyre, the work is recovered except for a fraction h we may write

$$\text{Work dissipated by wheel per revolution} = hp_0su.2\pi b \quad (1)$$

If v is the velocity of the driving vehicle, the number of revolutions per unit time is $v/2\pi b$, hence the rate of working per wheel is hvp_0su, and the rolling resistance is defined by the expression

$$R = hp_0su \quad (2)$$

Let the flattened zone of contact between tyre and plane have breadth e; then, if the distortion of the neighbouring portions of the tyre can be neglected, $e^2 \simeq 8bu$, and

$$p_0 = W/2s(2bu)^{\frac{1}{2}} \quad (3)$$

Equations (2) and (3) contain three unknowns, R, p_0 and u, so that we need a further relation between p_0 and u to solve for R. The assumption made is that in the narrow zone of contact between tyre and plane the elastic stresses are the same as if the whole surface of the tyre were acted upon by an external uniform normal stress p_0. This enables us to obtain a relation between p_0 and u by means of Lamé's well-known solution for the elastic stresses in a hollow cylinder submitted to external and internal pressures.[2] If these are p_0 and p_i respectively, we have the following solution for the stresses at a radius r:

$$\begin{aligned} \sigma_r &= (A/r^2) + 2C \\ \sigma_\theta &= -(A/r^2) + 2C \end{aligned} \quad (4)$$

where $$A = \frac{a^2b^2(p_0 - p_i)}{b^2 - a^2}, \quad 2C = \frac{p_ia^2 - p_0b^2}{b_2 - a^2}$$

The general expressions in polar co-ordinates for the displacements relating to a symmetrical stress distribution in a cylinder are[2]

$$\begin{aligned} u &= \frac{1}{E}\left[-\frac{(1+\nu)A}{r} + 2C(1-\nu)r\right] + H\sin\theta + K\cos\theta \\ v &= Fr + H\cos\theta - K\sin\theta \end{aligned} \quad (5)$$

where u is the radial and v the tangential displacement. E is Young's modulus for the tyre material and ν Poisson's ratio. A and C are defined by equations (4), and F, H and K are determined by the boundary conditions. Since the displacements are independent of θ, $H = K = 0$. Also, in this application of the theory there is a rigid core in place of the hollow of Lamé's solution so that $u = v = 0$ at $r = a$. Hence $F = 0$, so that the tangential displacement is everywhere zero, and also

$$0 = \frac{1}{E}\left[-\frac{(1+\nu)A}{a} + 2C(1-\nu)a\right] \quad (6)$$

Equation (6) can be solved for p_i, giving

$$p_i = 2p_0b^2/[a^2 + b^2 + \nu(b^2 - a^2)]$$

For rubber we may put $\nu = \frac{1}{2}$ and therefore

$$p_i = 4p_0b^2/(a^2 + 3b^2) \tag{7}$$

This value is substituted in equation (5), giving for $r = b$,

$$Eu = -\frac{3}{2}p_0(b - a)\left(\frac{ab + b^2}{a^2 + 3b^2}\right) \tag{8}$$

The negative sign enters because u is a compression, and no ambiguity will arise if it is omitted. Further, for a thin tyre on a large wheel, a and b are very nearly equal and $(b - a)$ is the tyre thickness t. Equation (8) may therefore be written

$$Eu = \frac{3}{4}p_0t \tag{9}$$

Eliminating p_0 and u between equations (2), (3) and (9),

$$R = (h/4.4)\ (W^4t/Esb^2)^{\frac{1}{3}} \tag{10}$$

EXPERIMENTAL RESULTS

A perusal of available records in which R was measured on a tyre-testing machine shows that the proportionality of rolling resistance to the cube root of tyre thickness is closely borne out (Fig. 1). Further, elimination of p_0 between

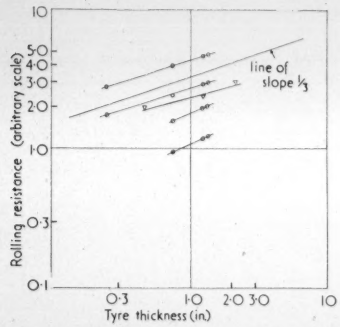


Fig. 1. Variation of rolling resistance with tyre thickness

equations (3) and (9) gives a relation between load and compression which should be applicable to static loading tests:

$$W = (8Es/3t)(2bu^3)^{\frac{1}{2}} \tag{11}$$

Here, again, the prediction that the slope of the line showing log W against log u should be 3/2 is very well corroborated over the range of load which is relevant to practice (Fig. 2).

As regards verification of equation (10) as a whole, experimental data for comparison were obtained from a tyre-testing rig in which a loaded wheel is driven by a drum of large diameter controlled by a d.c. motor, the power input to the motor being measured. Similar wheels of 30 in. diameter with three different thicknesses, 2·16 in., 1·26 in. and 0·47 in. were tested at two loads, 40 cwt and 71½ cwt, and at equivalent driving speeds of 10, 15, 20, 25 and 30 m.p.h. The tests can be regarded as only very rough since the rig is used normally for routine testing and is far from satisfactory for research purposes. The effective value of E, deduced from the compression test of Fig. 2 by means of equation (11), is 1900 lb/in.^{2}. Unfortunately, during the test (which had been carried out for another purpose), no

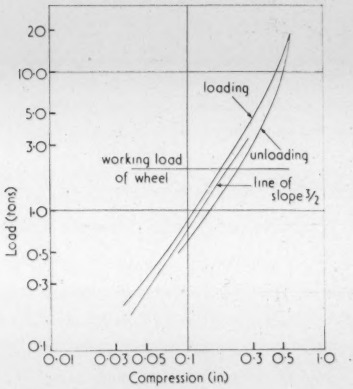


Fig. 2. Static load-compression test for wheel

Dimensions of wheel: diameter of rigid core = 27·4 in.; tyre thickness = 2·16 in.; tyre width = 4·75 in.

hysteresis loops had been traced relating to the loadings used in the rolling tests, and for the value of h it has been necessary to accept the manufacturers' quotation of 0·2, measured by the Dunlop pendulum.

Comparison of observed and calculated values of rolling resistance

Load (cwt)	*Tyre thickness (in.)*	*Rolling resistance (lb) Observed*	*Calculated*
40	2·16	38·2	34·8
	1·26	30·1	29·1
	0·47	25·1	21·0
71½	2·16	80·6	75·8
	1·26	54·4	63·3
	0·47	44·1	45·5

It can be seen from the table that observed and calculated values are roughly in agreement. That may be partly fortuitous, owing to uncertainty in the value of h. However, in view of the other evidence presented, and especially also in view of Tabor's validation of the fundamental physical assumption, it is possible to suggest that equation (10) does give a reasonably accurate value for rolling resistance.

ACKNOWLEDGEMENT

This paper is published by permission of the Chief Scientist, Ministry of Supply.

REFERENCES

(1) TABOR, D. *Phil. Mag.*, **43**, p. 1055 (1952).
(2) TIMOSHENKO, S. *Theory of Elasticity*, 1st ed., pp. 56–64 (New York: McGraw-Hill Book Co., Inc., 1934).

The growth of the liquid bridge in an electrical contact

By P. M. Davidson, B.A., Ph.D., Department of Physics, University College of Swansea

[*Paper received* 28 *April*, 1953]

This paper describes a mathematical procedure, the purpose of which is to calculate the progressive changes in the shape, size and temperature distribution of a molten contact-bridge, from the initial stage at which melting begins up to the final stage at which the bridge breaks.

When two contact electrodes are separating, the area of contact diminishes and the local temperature rises, until, at some isothermal surface in the contact region, the melting point is reached; with further separation the maximum temperature in the electrodes increases, so that the molten region grows from a surface to a volume, bridging the electrodes. Finally, the boiling point is reached at some isothermal surface in the bridge, which then breaks, leaving some of its metal on each electrode. The practical importance of the process is that, owing to the Thomson effect or some other cause, it produces a progressive transfer of metal from one electrode to the other. This is known as bridge transfer, to distinguish it from transfer due to electric discharges occurring either at make, or at break (after the bridge has broken). Recent experiments give a delicate method of measuring the contribution of bridge transfer to the total.[1] The theory of the bridge, on the other hand, is far from complete; even the valuable work of Lander and Germer[2] does not attempt to solve what is evidently a fundamental problem in the theory of the bridge, the detailed calculation of its shape and temperature distribution at successive stages in the separation, and in particular, of course, at the breaking stage.

THE GROWTH OF THE BRIDGE, FROM MELTING TO BOILING

For simplicity, we consider a mono-metallic contact pair, free from surface film; neglecting for the moment the Thomson effect, we will suppose that the initial melting is in the interface itself.

If we now make a slight separation of the electrodes, accompanied by an increase in contact potential, the molten region changes from a plane to a volume. We must consider the shape of the free liquid surface. If we neglect the electrodynamic forces and the variation of surface tension with temperature, the liquid is at rest and at uniform pressure, and with a uniform surface tension. The effect of gravity is negligible. Our calculations will make use of a theorem, given in text-books on surface tension, which states that under these conditions the possible surface shapes, symmetrical about an axis, are those whose intersections with a plane through the axis are traced out by a focus of a conic section rolled along the axis; and that according as the conic section is a hyperbola or an ellipse the surfaces are of two types:

(i) nodoids; those in the figures have p negative (by which we mean that the pressure in the liquid is less than the external atmospheric pressure); and

(ii) unduloids, of which the cylinder is a particular case, the general unduloid differing from the cylinder by having a wave-like surface, causing the width to oscillate between a maximum and a minimum value. They have positive p.

The intermediate case ($p = 0$) between the nodoid and unduloid form is the catenoid (the section being a catenary, obtained by rolling a parabola).

Starting then from the melting stage, and moving the electrodes slightly apart, we can now calculate what will happen, if we know the circuit and the properties of the metal. Suppose, for example, that the external e.m.f. is large enough to keep the current practically constant as the bridge is extended. We have to draw, as the free surface of the liquid, one of the possible types of curve described above, such that the total volume of metal is unaltered (after correcting for thermal changes in volume) and such that the (unchanged) current through the bridge raises the point where the curve intersects the electrodes to the melting point. This calculation, giving the shape and size of the bridge, can be made, and proceeding in this way, we can trace the successive changes as the electrodes are progressively separated. These changes will depend very much on the properties of the particular metal (e.g. on the variation of resistance with temperature and the thermal changes in volume*) and also on the external circuit, which determines whether there is a small or large change in the current as the bridge potential changes from the melting to the boiling voltage. Contrasting cases are sketched in the figures. In set *A*, Fig. 1, we have a case where, as the electrodes separate, the solid-liquid junction moves continuously away from the axis, and the curves are all nodoids. Set *B* indicates how these diagrams would be modified by a considerable Thomson effect, or other cause of asymmetry. Below the last member of set *B* is drawn the same diagram as seen from the side. The breaking of the bridge by boiling results in a gain of metal by the right hand electrode. Set *C*, Fig. 2, is a set in which, during the later stages of the separation, the solid-liquid junction moves towards the axis; the liquid surface, initially a nodoid, has become a catenoid in the second diagram; it has then become a cylinder (at a stage not shown) and finally, in the third diagram, an unduloid. Set *D* indicates how set *C* would be modified by a considerable Thomson effect, and below the last member of set *D* is shown the same diagram viewed from the side; if at this stage the boiling voltage is attained, the breaking of the bridge at the hottest section, which is not far from the narrowest part of the bridge, leaves a pip on the right hand electrode and a crater in the other.

It is satisfactory to note that in the two diagrams of bridges seen from the side, the shapes are typical of the concave and cup-shaped bridges observed experimentally with a microscope.[3]

There are, however, some matters which must be considered more fully, in individual cases, to decide whether the above procedure would give a good approximation. We have assumed that the surface tension was constant over the surface; but actually a metal surface, if uncontaminated, has a smaller surface tension in the hot region, near the centre, than in the colder regions near the side. This variation of surface tension along the length of the bridge will produce something more than a modification of shape; a liquid with a

* For example, in the actual process of melting some metals increase in specific resistance by factors of the order of 3, whereas others diminish; some expand by ten per cent or more, while others contract.

non-uniform surface tension cannot be in equilibrium and at rest. A circulating motion will be set up in the bridge, and it is the viscous forces in the surface region which balance the variation in surface tension. If, on the other hand, the surface is contaminated, matters are different. The surface tension depends on the thickness of the contaminating film; the beginning of the circulating motion will redistribute the thickness of the contaminating film until, in spite of the variation in temperature, the surface tension is uniform over the surface and the bridge is at rest. Thus bridges may be divided into two classes—those which have a pure surface, and which, on that account alone, can never be free from internal motion; and those which have a contaminated surface, with what we may call an adjustable surface tension, not determined by the temperature.

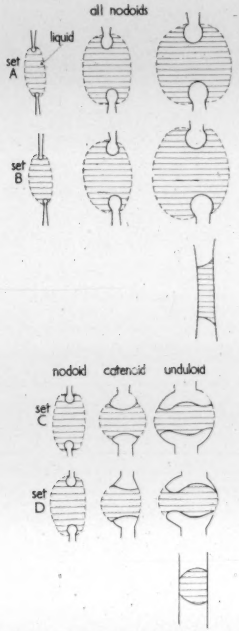


Fig. 1

Fig. 2

Figs. 1 and 2. Types of bridge development

Even in bridges of the latter type the surface shape will always be more or less modified by the electrodynamic forces, which we have neglected. It must be remembered that there is a magnetic field in the bridge, the lines of force consisting of circles round the axis. Each element of liquid metal is conveying a current, and (since the current is flowing in a magnetic field) is acted on by a mechanical force at right angles to the current and in a plane through the axis. Unless the bridge is cylindrical these electrodynamic forces are not of the simple type which can be entirely balanced by a redistribution of pressure in the bridge (accompanied by a modification of the surface shape); they give rise also to circulating motions in the fluid, though not necessarily of its outer surface, if it is contaminated. Thus the electrodynamic forces are balanced partly by a pressure distribution and partly by viscous forces. It is only in very small bridges that the electrodynamic forces are negligible compared with surface tension forces and produce only negligible changes in the shape of the bridge.

Finally, the question of instability has to be considered and only one type of instability will be mentioned here. Suppose we have a metal for which, in a contact containing a liquid bridge, the current i, contact potential V, and apparent electrode separation l are connected by a relation of the form

$$Vi = f(l) \quad \text{where } f(l) \text{ increases with } l. \tag{1}$$

Changes in the apparent electrode separation can be measured, for example, by a screw which moves one of the electrodes; owing to local heating, these changes are not usually the same as the changes in the local interfacial separation at the contact region.

For our purpose it does not matter what the function $f(l)$ is, provided it increases with l. [A relation approximately of this form was found by Betteridge and Laird[3] and by Fairweather[4] for the metals (platinum and certain others) which they employed.] Suppose, now, that we consider a simple circuit consisting of a battery of e.m.f. E and two resistances R and r in series. The circuit will also, of course, have some inductance. If the potential across r is V we have, for a steady current, $Ri = E - V$, and thus, multiplying through by V, we have $RVi = EV - V^2$. Plotting RVi against V we obtain a parabola on which RVi is zero at $V = 0$ and $V = E$, and has a maximum $\frac{1}{4}E^2$ at $V = \frac{1}{2}E$. If both R and r were ordinary variable resistances, we could, by varying either of them, obtain the whole curve from $V = 0$ to $V = E$: but this is not so if r is a contact with a liquid bridge satisfying relation (1). For $\frac{1}{2}E$ must be either greater or less than the boiling voltage. Suppose first that it is greater, and that we slowly increase l, and hence Vi, at constant R (or slowly increase R at constant l) from a value of RVi at which the smaller of the two V's given by the parabola is less than the boiling voltage. (It is readily shown that the smaller of the two V's is a stable state in this circuit: the larger V, being greater than the boiling voltage, is unobtainable.) Thus as we slowly increase l (or R) we pass up the parabola until V attains the boiling voltage, and the bridge then breaks: this is, of course, the usual phenomenon of breaking at the boiling voltage. But now suppose $\frac{1}{2}E$ is less than the boiling voltage, though still greater than the melting voltage. As we slowly increase l at constant R (or R at constant l) we again pass up the parabola on the side with the smaller V's (at a given RVi the smaller value of V is a stable state in this circuit, whereas the larger value of V, even if less than the boiling voltage, is an unstable state and cannot be obtained). Having thus increased l (or R) until the maximum of the parabola (less than the boiling voltage) has been reached, suppose we increase l (or R) very slightly further. The relation (1) requires that RVi shall increase; but there is no such point on the parabola. To trace in detail what will happen we must take account of the inductance; we then find that the infinitesimal increase in l (or R) causes the contact potential V to increase "explosively." In the course of this potential rise the bridge will boil and break. Strictly speaking, the break is still due to boiling, but the experimenter will say that the bridge broke, not at the normal boiling voltage but at the smaller value $\frac{1}{2}E$. Thus, for example, if platinum/iridium has a boiling voltage of 1·6, and if it accurately obeys the law (1), then in simple circuits

of the type described, the bridge will break at the normal boiling voltage if E is greater than 3·2 V, but if E is only 2 V it will not be possible to obtain stable bridges with V greater than 1 V. There seems to be some experimental evidence that in 2 V circuits, bridges of platinum[5] and some other metals break at less than the normal boiling voltage; and such agreement is the most that can be expected; for it must be remembered that our result depends on assuming that the relation (1) holds accurately in the region of the breaking voltage, whereas actually relation (1) is only a rough empirical law, not likely to be satisfied accurately by any metal. In metals for which it is not even approximately satisfied there is, of course, no reason to expect that there will be any instability of the type considered.

Throughout this paper we have thought of the electrode faces as being separated extremely slowly (as they often are, in experiments). The calculations would, however, be only slightly modified by velocities such as arise in practical contacts (except that the occurrence of instability, if any, would, of course, be less sharply observable).

ACKNOWLEDGEMENTS

This paper forms part of a programme of research (on the fundamental processes in electrical contacts), being carried out under the direction of Professor Llewellyn Jones, in the Physics Department of the University College of Swansea.

REFERENCES

(1) WARHAM, J. *Proc. Instn Elect. Engrs*, **100**, I, p. 163 (1953).
(2) LANDER, J. J., and GERMER, L. E. *J. Appl. Phys.*, **19**, p. 910 (1948).
(3) LLEWELLYN JONES, F. *Fundamental Processes in Electrical Contact Phenomena*, H.M.S.O. Special Report No. 24 (May, 1953).
BETTERIDGE, W., and LAIRD, J. H. *J. Instn Elec. Engrs*, **82**, p. 623 (1928);
PAETOW, H. *Elektrotech. Z.* [*ETZ*], **70**, p. 227 (1949).
(4) FAIRWEATHER, A. *J. Instn Elec. Engrs*, **304**, p. 92 (1945).
(5) WARHAM, J. Private communication; and experiments by JONES, R. H., in the Physics Department at Swansea.

Automatic electron trajectory plotting using the electrolytic tank analogue*

By B. O. BAKER, B.Sc.(Eng.), A.C.G.I., A.M.I.E.E., Research Laboratories, The General Electric Company, Ltd., Wembley

[*Paper first received* 21 *July, and in final form* 23 *September*, 1953]

Equipment is described for the automatic plotting of electron trajectories with the electrolytic tank analogue. It has been found that the results in a known system are within 2% of the correct solution. The method is applicable to electron optical problems in two or three dimensions. The tank has also been used to solve a variety of problems involving Laplace's equation to an accuracy of better than 1%.

INTRODUCTION

The solution of many problems in electron optics requires a knowledge of the paths of electrons in the system. For complex systems which do not readily lend themselves to analytical methods, the electrolytic tank analogue can be used. Originally the electrostatic field was plotted in the tank, and the paths of electrons were determined graphically by successive applications of Snell's law. This is very tedious, and Gabor[1] developed a method of plotting trajectories manually. Langmuir[2] produced an automatic system for solving the problem, and the present equipment has been developed from the suggestions in his paper. Attention has so far only been directed to electron trajectory plotting in electrostatic fields where the influence of space charge can be neglected or calculated, but it is hoped at a later date to provide additional equipment to take account of space charge.

GENERAL CONSTRUCTION

The general arrangement of the tank and equipment is shown in Fig. 1.

Tank. A large tank 3 ft × 5 ft was chosen in order to achieve reasonable accuracy without the need for elaborate precautions to avoid errors due to such factors as polarization, misalignment of the electrodes and edge effects from the sides of the tank. This was constructed in perspex, supported in an iron framework on three jacks, to provide means of tilting the tank through 10° about an axis parallel to either the long or short sides. By this means it is possible to solve problems in three dimensions having axial symmetry, using simple electrodes and a shallow tank. A flat bottom is obtained with a sheet of glass resting in the tank. For some applications a deep tank is necessary, and provision has been made to fit a tank 20 in. deep when required. This enables

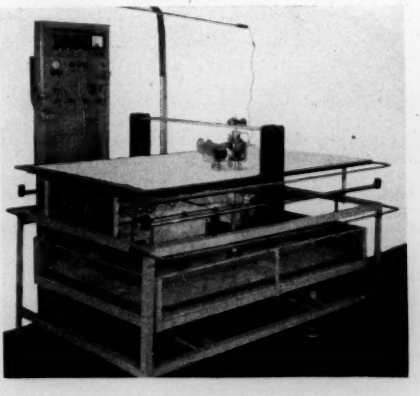


Fig. 1. General arrangement of tank and equipment

* G.E.C. Communication No. 541.

trajectories to be plotted through the axis of symmetry, which cannot be done with the tilted tank.

Measurements in the tank are made with probes just in the surface of the electrolyte, and it is necessary to be able to move these over the whole surface of the electrolyte without altering appreciably their depth of immersion. This is achieved with a gantry running on rails parallel to the long sides of the tank, accurately adjusted to be straight and horizontal. The gantry carries two pairs of transverse rails, one above and one below the plotting table. The lower pair of rails support a probe carrier, and the upper pair a plotting carrier and these are linked by wires so that the motion of one is accurately reproduced by the other without any backlash. The overall accuracy of this part of the system was checked with a dial gauge supported from the probe carrier and measuring the flatness of the glass plate in the tank. With the tank accurately horizontal the spacing between the probe carrier and the bottom of the tank is constant within $\pm 0{\cdot}005$ in.

Choice of electrolyte. Electrodes and probes. In a tank of this size one of the most important factors governing the choice of electrolyte is that it must be homogeneous and capable of easy replacement to allow for evaporation. The tap water available, with a specific conductivity about 1 millimho per centimetre, has been found suitable. McDonald[3] and others have also found tap water to be a reasonable electrolyte. Polished brass electrodes have been used, and graphite can be painted on the brass to reduce polarization, but this has not normally been required. The probes used for measurement in the electrolyte are of platinum coated with a sponge-like layer of platinum black[4] to increase its surface area. Errors due to polarization[5,6] can be more serious at the probes than at the electrodes. The operating frequency of 1000 c/s was chosen after preliminary experiments at various frequencies. It has been found in practice that repeatable results can be obtained from day to day, and these do not appear to be affected by the dust which inevitably collects on the electrolyte surface during the solution of any particular problem.

AUTOMATIC TRAJECTORY PLOTTING

Various systems for plotting electron trajectories automatically have been proposed.[1,2,7] The system used is a development of Langmuir's method,[2] whereby the instantaneous radius of curvature of an electron path is determined.

The path of an electron at a particular point in an electrostatic field is determined by the potential V at that point and the potential gradient ϵ_n normal to the path. From the equations of motion of an electron it can be shown that the instantaneous radius of curvature is given by

$$\rho = \frac{2V}{\epsilon_n} \tag{1}$$

The value of ρ is derived by measuring ϵ_n and V by means of two probes in the surface of the electrolyte with a fixed spacing between them.

The method of determining the ratio ϵ_n/V is shown in Fig. 2. The sliding contact of the potentiometer is automatically adjusted by a control unit described later, to give zero output from the null amplifier. The sliding contact of the potentiometer is linked by selsyns to a steering trolley shown schematically in Fig. 3, which drives the gantry, and also plots the electron path. The radius of curvature of the path is continuously adjusted automatically as the potentiometer finds its balance position to measure ϵ_n/V. This is done by means of lead screws C, C', which are driven by the selsyns linking the potentiometer and the trolley. The lead screws carry pins D and D' which traverse the whole length $2p$ of the lead screws in opposite directions as the potentiometer slider traverses its whole length $2g$. The arm AOA' is spring loaded on to one or other of the pins, according to whether the trolley is steering to the left or the right.

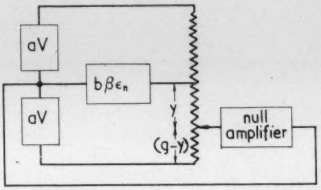


Fig. 2. Potentiometer circuit for determining ϵ_n/V

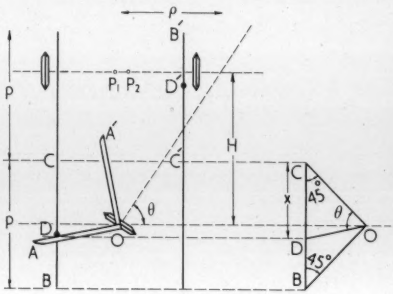


Fig. 3. Schematic diagram of trolley

$$\frac{x}{p-x} = \tan\theta = \frac{H}{\rho}$$

With a = gain of V amplifier,
b = gain of ϵ_n amplifier,
β = separation of probes,

the condition of balance from Fig. 2 is

$$\frac{\epsilon_n}{V} = \frac{a}{b\beta}\frac{y}{(g-y)} \tag{2}$$

and from Fig. 3

$$\frac{2}{\rho} = \frac{2}{H}\frac{x}{(\rho - x)} \tag{3}$$

Providing the potentiometer and the lead screws are suitably linked,

$$\frac{y}{g-y} = \frac{x}{\rho - x} \tag{4}$$

The condition for correct trajectory plotting then reduces to

$$\frac{2}{H} = \frac{a}{b\beta} \tag{5}$$

This can be achieved by adjusting the gains a and b of the two amplifiers.

CONTROL EQUIPMENT

1000 c/s supplies. The 1000 c/s supply for the electrodes in the tank is derived from a conventional Colpitt's oscillator and 30 W amplifier feeding 20 V to a distribution panel. Provision is made to apply a small voltage of the appropriate phase to the cathode electrode in the tank if it is required to take account of finite velocities of electron emission. Voltages applied to the various electrodes are measured by a null bridge method to avoid errors due to loading the system with the measuring instrument.

Probe mounting. The probes consist of two platinized platinum wires 0·6 mm diameter, 2·0 mm apart, coated with glass to within 2·0 mm of the tip. These probes are mounted with the glassed portion in the surface of the electrolyte, so that the potentials are measured just below the surface to eliminate surface tension effects. Electrical contact to the probes is made through watch springs, arranged in opposition so that they do not produce any appreciable torque. Sander, Oatley and Yates[7] have considered the effect of inserting probes of finite size into the tank. With the probes described surrounded by screening at cathode potential. Some phase adjustment is provided in the V amplifier so that when the V and ϵ_n signals are balanced against each other as shown in Fig. 2; any slight difference in phase can be eliminated. The output of the V amplifier is accurately balanced with respect to earth with a pre-set control.

Power supplies to both amplifiers are derived from stabilized power packs of conventional design. The negative feedback in both amplifiers ensures high stability of gain which is essential for accurate trajectory plotting.

Null amplifier. The out-of-balance voltage from the potentiometer of Fig. 2 is fed into a null amplifier. This is a four-stage amplifier with an overall voltage gain of 11 000, sharply tuned to 1000 c/s. Automatic volume control is incorporated to limit the output of this amplifier to about 7 V, which is necessary for the correct operation of the control unit. A cathode-ray tube monitor is used to observe the null amplifier output, and phones can be plugged in if required.

Control unit. The potentiometer is a helix of forty turns of Eureka wire wound on an insulating former 2 in. in diameter. The arm carrying the sliding contact to the potentiometer is motor driven. As the sliding contact passes through the balance point, the phase of the out-

R_1 = 1 MΩ
R_2 = 22 kΩ
R_3 = 4·7 kΩ
R_4 = 330 Ω
R_5 = 470 Ω
R_6 = 470 Ω
R_7 = 330 kΩ
R_8 = 22 kΩ
R_9 = 330 Ω
R_{10} = 47 kΩ
R_{11} = 3·3 kΩ wire
R_{12} = 1 MΩ
R_{13} = 330 Ω
R_{14} = 100 Ω
R_{15} = 1 MΩ
R_{16} = 150 Ω
R_{17} = 1 MΩ
R_{18} = 2·2 MΩ
R_{19} = 330 Ω
R_{20} = 10 kΩ
R_{21} = 10 kΩ
R_{22} = 100 kΩ
R_{23} = 4·7 kΩ
R_{24} = 10 kΩ
R_{25} = 0·75 Ω

C_1 = 0·5 μF
C_2 = 0·05 μF
C_3 = 0·01 μF
C_4 = 0·5 μF
C_5 = 100 pF
C_6 = 0·01 μF
C_7 = 20 μF
C_8 = 0·003 μF
C_9 = 1 μF
C_{10} = 0·1 μF
C_{11} = 0·1 μF
C_{12} = 0·02 μF
C_{13} = 0·002 μF
C_{14} = 1 μF

L_1 = 200 H
L_2 = 20 H
L_3 = 10 mH
L_4 = 20 H

Resistors marked *s* are stoppers

All valves are type Z77

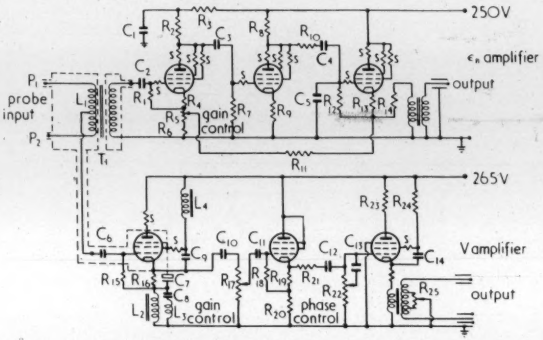


Fig. 4. Circuits of V and ϵ_n amplifiers

here, the error due to interaction is negligible if the angle of incidence between the probes and the equipotentials does not exceed 60°.

In order to measure the gradient ϵ_n it is necessary to keep the probes perpendicular to the direction of motion, and this is done by linking the probe mounting to the steering trolley by means of selsyns. The position of the probes is indicated by P_1 and P_2 in Fig. 3. Complete screening of the leads to the probes is necessary to avoid stray capacity effects. This is described in the next section.

The probes are surrounded by a guard ring, so that if the structure accidentally strikes an electrode in the tank, a relay operates to switch off the trolley driving motor, avoiding damage to the probes.

V and ϵ_n amplifiers. The circuits of the V and ϵ_n amplifiers are shown in Fig. 4. If the screening of the probes P_1 and P_2 were connected to earth, there would be a spurious potential difference between them due to an unbalanced stray capacity. This can be reduced if the potential of the screening is very near that of the probes. To achieve this, the centre tap of the probe transformer T_1 is connected to the grid of a cathode follower with a degeneration of about 1000, and the screening to its cathode. The primary winding of T_1 is also completely

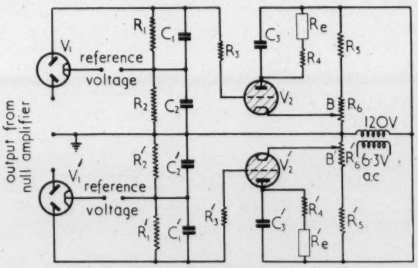


Fig. 5. Circuit diagram of control unit

R_1, R_1' = 100 kΩ
R_2, R_2' = 100 kΩ
R_3, R_3' = 220 kΩ
R_4, R_4' = 33 kΩ
R_5, R_5' = 100 kΩ
R_6, R_6' = 10 kΩ
R_e, R_e' = 2 mA relay
C_1, C_1' = 0·05 μF
C_2, C_2' = 0·05 μF
C_3, C_3' = 2 μF
V_1, V_1' = D63
V_2, V_2' = GT1C

of-balance voltage changes by 180°. This out-of-balance voltage is fed into the control unit as shown in Fig. 5, and its phase is compared with that of a reference voltage. If they are in phase, one of the thyratrons V_2 and V_2' strikes, and if they are out of phase the other strikes. High speed relays in the thyratron anode circuits operate contacts which switch the potentiometer motor in such a way as to restore balance. In practice the inertia of the mechanical system causes the potentiometer to overshoot, and hunting takes place. The amplitude of this is too small to be observed on the trace of the trajectory.

Trolley. Plotting of the electron trajectory is done by means of a trolley linked to the gantry as shown in Fig. 1. The trolley is electrically driven, and is steered by means of the lead screws (Fig. 6). A spring loaded pencil is mounted

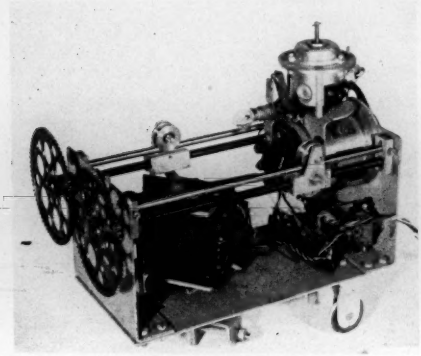


Fig. 6. Trolley used for plotting

on the axis of the trolley mid-way between the two front wheels, so that it plots the path of the point mid-way between the two probes.

To facilitate accurate steering, the wheels of the trolley are shaped to give a small area of contact with the plotting table, the friction between the wheels and the paper being sufficient to prevent slipping. This can be demonstrated by setting the trolley to a fixed radius of curvature, when it will plot very accurate circles while driving the gantry and probe structure. Care had to be taken in the construction of the trolley that the front wheels were accurately parallel, to ensure correct tracking.

The radius of curvature of the path which can be plotted is only limited by the finite separation of the probes. There is, however, a limit to the speed of steering, determined by the torque available from the steering selsyns. For this reason a two-speed switch is provided for the driving motor, so that it can be slowed down when a sharp change of curvature of trajectory is being plotted. The speeds of plotting are about 4 and 0·7 mm/sec. An average trajectory takes only about 3 min to plot.

METHOD OF OPERATION

In order to make the necessary adjustments for trajectory plotting, it is convenient to set up a planar diode in the tank. This consists of two parallel flat brass plates, bounded at each end by insulating perspex plates perpendicular to the conductors.

The first adjustment is to synchronize the trolley steering with the potentiometer, and this is done by feeding only the V signal to the potentiometer. Under these conditions the system hunts about the mid-point of the potentiometer providing the V output is correctly balanced with respect to earth. The trolley is then adjusted to steer in a straight line. The probes are aligned correctly by setting them to give zero output when the trolley is travelling perpendicular to the electrodes of the diode.

In order to adjust the gains of the V and ϵ_n amplifiers to satisfy equation (5), it is convenient to set the trolley in the centre of the diode steering parallel to the electrodes. Under these conditions, if the electrode spacing is d and the voltage between them V',

$$V = \tfrac{1}{2}V' \quad \text{and} \quad \epsilon_n = V'/d$$

Hence from equation (1): $\rho = d$.

The potentiometer contact is now adjusted so that the trolley will steer a circle whose radius is equal to the spacing between the diode electrodes. The gain and phase of the ϵ_n and V amplifiers are then adjusted to give zero output from the null amplifier. In order to avoid the necessity of repeating this operation during the solution of any particular problem, which might last for several days, provision has been made for switching in a calibrating voltage in place of the output from the probes.

The overall accuracy of trajectory plotting can be determined with the planar diode. If trajectories are plotted as in Fig. 7, starting parallel to the electrodes, the path is parabolic, and can readily be calculated. The overall accuracy determined in this way is within 2%. These trajectories can be repeated within 2 mm on a 32 cm track. An interesting check has been made with the tank tilted for three dimensional problems with axial symmetry. If the same planar diode is placed with the electrodes perpendicular to the shore of the electrolyte, it represents two infinite parallel flat planes, and the trajectories as in Fig. 7 can be repeated with the same

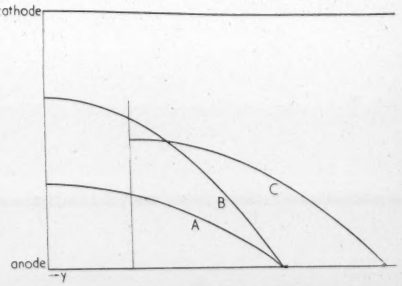


Fig. 7. Electron trajectories in a planar diode

trajectory	*y* in. *measured*	*y* in. *calculated*
A	10·28	10·40
B	10·32	10·40
C	11·20	11·06

degree of accuracy. In this case the electrolyte depth varies along the trajectory, but this does not introduce appreciable errors. In the case of finding the focal length of cylindrical

systems for paraxial rays, due to the small angle the ray makes with the axis, small errors in the measurement of ϵ_n produce large errors in the focal length. The error in determining the focal length from paraxial trajectories is likely to exceed 2%.

PRACTICAL APPLICATION

The tank has been used successfully in the solution of several problems involving deflexion or focusing of electron beams. An example is the design of an electron gun with a rectangular cathode. This was a problem in which an approximate solution could be obtained neglecting space charge. A rectangular Pierce type gun had been designed theoretically, and it was desired to check the results in the electrolytic tank before constructing the prototype, particularly to find out how critical were the electrode positions. The results obtained in Fig. 8 showed good agreement with the theoretical results neglecting space charge. It was calculated that, if under these conditions the electron beam focused near the opening to the first anode, the effect of space charge would be to cause the beam to pass through the subsequent electrodes without interception.

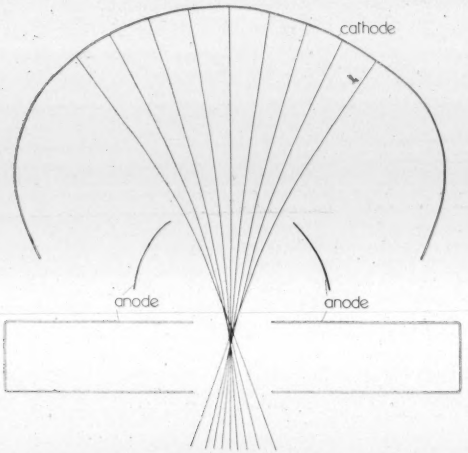


Fig. 8. Electron trajectories in rectangular Pierce type gun

EQUIPOTENTIAL PLOTTING

Provision has been made for field plotting, as this greatly increases the applications for which the tank can be used. The method is conventional, and makes use of the equipment which has already been described.

Field plotting has been used in the design of high voltage cable terminations, using a wedge shaped model in which the effect of changes in dielectric constant can be simulated by abrupt changes in electrolyte conductivity, in the manner described by McDonald,[3] using moulded wax models.

No attempt has been made to plot the equipotential automatically, as the complication involved is not justified. Manual plotting is very rapid, using either the cathode-ray tube monitor or phones as the null detector. The accuracy of field plotting is better than 2% with wax models simulating various dielectrics, and better than 1% with single dielectrics.

DISCUSSION

The accuracy achieved with the present equipment should be adequate for most practical purposes, and the method described for trajectory plotting is very rapid. The equipment used is relatively simple, which is a great advantage where a digital computor[7] is not available. The only limit to the radius of curvature of the trajectories is set by the finite size of the probes, and it is not expected that practical problems will be affected by these considerations.

So far in the work no account has been taken of the effect of space charge. A method has been proposed[8] for modifying the bottom of the tank to allow for space charge. A more direct analogy would be to inject currents into the tank from a lattice of probes, modifying the general current flow. It is proposed to investigate this method, which would probably be more rapid and easily applied to a variety of problems.

REFERENCES

(1) Gabor, D. *Nature* [*London*], **139,** p. 373 (1937).
(2) Langmuir, D. B. *R.C.A. Rev.*, **11,** pp. 143–154 (1950).
(3) McDonald, D. *Proc. Instn Elect. Engrs*, Pt 2, **100,** pp. 145–166, 176–183 (1953).
(4) Jones, G., and Bollinger, D. M. *J. Amer. Chem. Soc.*, **57,** pp. 280–284 (1935).
(5) Jones, G., and Christian, S. M. *J. Amer. Chem. Soc.*, **57,** pp. 272–280 (1935).
(6) Cherry, E. C., Boothroyd, A. R., and Makar, R. *Proc. Instn Elect. Engrs*, Pt 1, **96** (1949).
(7) Sander, K. F., Oatley, C. W., and Yates, J. G. *Proc. Instn Elect. Engrs*, Pt 3, **99,** pp. 169–179 (1952).
(8) Goudet, G., and Musson-Genon, R. *J. Phys. Radium*, **6,** pp. 185–195 (1945).

Correspondence

A simple bridge method of measurement of the time constant of exponential decays

In a letter published in the November 1953 issue of your *Journal*, Mr. C. M. Burrell describes a simple bridge method of measurement of the time constant of exponential decays. He shows that if R' is zero, there is no current in it during the decay of the oscillation if $CR = 1/\alpha$, since the currents through R and C balance. The assumption that these currents balance even if R' is finite is, however, erroneous, because the initial charge on C and current through R are modified. In fact the output voltage will decay from its value $V_0R'/(R + R')$ at the end of the pulse with a time constant $RR'C/(R + R')$. The conclusion is thus approximately correct only if $R' \ll R$. Alternatively, it could be made exactly correct by removing a short circuit across R' at the end of the pulse.

H. H. Wills Physical Laboratory, D. F. GIBBS
University of Bristol.

Mr. Gibbs' analysis is quite correct as far as it goes and I am grateful to him for pointing out this error. However, it will be noted that, so long as $R' \ll R$, the error due to the transient he mentions decreases rapidly with time. Thus if $R'/R = 0{\cdot}05$, the error voltage will have decreased to approximately $0{\cdot}005\%$ of its original value after a time $t = 1/2CR$, whereas the required error signal will still be about 60% of its original value, giving a ratio of approximately 10000 : 1 between the required and unwanted signals. Hence the restrictions (i) $R' \ll R$ and (ii) observation beginning after $t = 1/2CR$ (say) are sufficient to ensure negligible error. These restrictions were in fact observed, for other reasons, in the experiments carried out using this circuit.

Radar Research Establishment, C. M. BURRELL
Malvern, Worcs.

The measurement of spectral distributions of luminescent materials above room temperature

Previous investigations* of the spectral distribution of luminescent materials above room temperature have involved the use of somewhat cumbersome apparatus. This may be overcome by using a conducting film as heating element.

As shown in Fig. 1, a silver-activated zinc sulphide phosphor, sandwiched between two glass slides, was placed

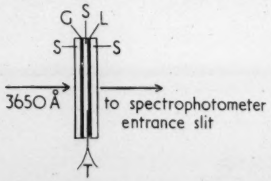


Fig. 1. Arrangement of phosphor and conducting film for 3650 Å excitation

S, glass slide;
G, gold film;
L, phosphor layer;
T, thermocouple.

in contact with a gold film mounted on a glass slide. Light of wavelength 3650 Å, obtained by passing the radiation from a high-pressure mercury vapour lamp through a Woods glass filter, was incident on the phosphor through the gold film. The spectral distribution of the luminescence was measured by means of a Unicam S.P.500 spectrophotometer.

* NICOLL, F. H. *J. Opt. Soc. Amer.*, **38**, p. 817 (1948).

The phosphor temperature was measured with a copper-constantan thermocouple, in contact with it; the temperature could be accurately controlled by adjusting the current through the conducting film. The spectral distribution of the luminescence of the gold film in the range of the phosphor luminescence was determined in a separate experiment and found to be negligible. The effect of any selective reflexion from the gold film was also negligible, since room temperature spectral distributions measured with and without the film were identical. Spectral distributions of the phosphor at 20°, 60° and 100° C are shown in Fig. 2; the significance of

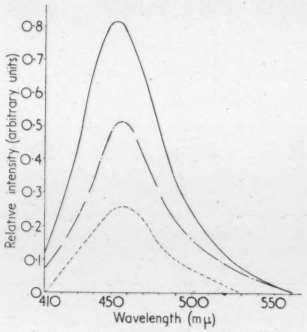


Fig. 2. Spectral distribution curves for silver activated zinc sulphide phosphor
——— = 20° C; — — — = 60° C; - - - - - = 100° C.

these results for luminescence theory will be discussed in a future publication.

The method is also applicable to excitation by wavelengths not transmitted by the conducting film; it is then, of course,

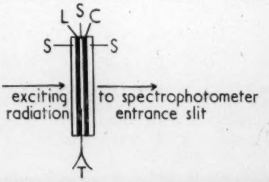


Fig. 3. Suggested arrangement of phosphor and conducting film for excitation by wavelengths not transmitted by the film

S, mounting slide;
L, phosphor layer;
C, conducting film;
T, thermocouple.

necessary to place the phosphor slide on the other side of the conducting film, as indicated in Fig. 3, and to allow for any non-uniformity in the spectral transmission curve of the film.

The authors are most grateful to Mr. J. S. Preston and Mr. E. J. Gillham of the National Physical Laboratory for the gift of the conducting film and to Dr. S. Rothschild of the Mullard Radio Valve Co. for the phosphor. They wish to offer their best thanks to Dr. J. F. J. Dippy of the Chemistry Department for permission to use the spectrophotometer and to Dr. A. M. James for his help in adapting the instrument for this investigation.

Physics Department, A. LANGRIDGE
Chelsea Polytechnic, MARY P. LORD
London, S.W.3.

New books

Methods of theoretical physics (Parts 1 and 2). By PHILIP M. MORSE and HERMAN FESHBACH. (London: McGraw-Hill Publishing Co., Ltd., 1953). Part one: Pp. xxii + 997 + xl. Part two: Pp. xviii + 999 to 1978. Price £6 each part.

The present reviewer has found it difficult to read this book because it has been borrowed so often by his colleagues. This is a tribute to its usefulness. In fact, this book and Harold and Bertha Jeffreys' *Methods of mathematical physics* are as far as he knows the only up-to-date sources in English to which a research worker in theoretical or mathematical physics can go to learn the technique of his subject, whether he wishes to spend a year learning it from the beginning or whether he wants to look up some method of which he was previously unaware.

Contrasting the two, the book by Morse and Feshbach is nearly three times the length, and it lacks the mathematical eloquence of the earlier book, giving the impression that it is written by and for the modern theoretical physicist in a hurry, who, living in the competitive and urgent atmosphere of American physics today, wants to get a physically acceptable answer to a problem before the next man, and cares not at all for the rigour of an older generation. Its emphasis, too, is different; whereas the background to the older book is astronomy and geophysics, Professor Morse and his colleague have been brought up on such problems as quantum mechanics, acoustics, radar and the diffusion of neutrons in piles. As the authors say, the book is about fields rather than particles, wave equations rather than orbits, and boundary conditions, Green's functions and characteristic values play a large part in it. Fluid mechanics is treated rather briefly; it is not the main centre of interest. On the other hand, a hundred pages are devoted to solutions of Schroedinger's equation. The book, however, very properly does not discuss modern and controversial matters, such as the theory of nuclear forces.

The book will be invaluable as a work of reference to research workers in all branches of theoretical physics which have to do with wave and diffusion equations. For a student beginning the subject it should be useful too, if he is not intimidated by its formidable length. There are examples for such a student to work out, which will help him, and which indeed will help anyone arranging a course of instruction on the subject. The authors say that, with some major interruptions, the book has taken them sixteen years to write. They are to be congratulated on finishing it.

N. F. MOTT

Complex variable theory and transform calculus (2nd edition). By N. W. MCLACHLAN. (London: Cambridge University Press.) Pp. xi + 388. Price 55*s*.

This book was first published in 1939 under the title *Complex variable and operational calculus*. It has now been revised, some old matter having been removed and some modern topics introduced. Part I which deals with the theory of the complex variable has been rewritten to meet the criticism of "certain pure mathematicians (and physicists!)" who in reviewing the first edition of the book commented on the inadequacy of the degree of rigour achieved. The author's claim that the treatment is now rigorous enough for all but the pure mathematician (and the book is not for him) must be admitted; indeed, this first part is a clear and readable account of contour integration and the associated complex variable theory.

Part II deals with the theory of transform calculus, and now includes Fourier transforms and frequency spectra. Part III is devoted to technical applications covering a wide field, namely electrical circuits, vibrational systems, aeroplane dynamics, radio and television receivers, as well as problems involving partial differential equations associated with transmission lines, wave filters and similar topics. This part should be very valuable to the technologist to whom the book is mainly addressed. There is an excellent list of references.

J. TOPPING

Progress in nuclear physics, Volume 3. Editor: O. R. FRISCH, O.B.E., F.R.S. (London: Pergamon Press Ltd.) Pp. vii + 279. Price 63*s*.

Professor Frisch's series is now sufficiently well known to need no introduction, and the new volume follows the same lines as the first two. It contains nine articles by active and established specialists, intended mainly for physicists working in related fields. M. Snowden writes on the diffusion cloud chamber; D. West on energy measurements with proportional counters; R. J. Blin-Stoyle, M. A. Grace and H. Halban on oriented nuclear systems; J. V. Jelley on Čerenkov radiation; M. Deutsch on annihilation of positrons; F. C. Champion on solid conduction counters; R. Huby on stripping reactions; P. C. Thonemann on the production of intense ion beams; and H. S. W. Massey on the collision of deuterons with nucleons.

Four of the articles deal with experimental results and the appropriate theory, the emphasis being mainly theoretical in the articles by Professor Massey and Dr. Huby, and the other five articles are concerned with tools of nuclear research. The excellent standards set in the earlier volumes have been well maintained, and again there is practically no overlapping with previous articles in the series, which has now covered considerable ground. It is noteworthy that each of the nine lists of references, appended to the articles, includes work published in 1953.

T. G. PICKAVANCE

Bessel functions and formulae. Compiled by W. G. BICKLEY. (London: Cambridge University Press.) Pp. xxx–xl. Price 3*s*. 6*d*.

This booklet is a reprint of part of the introduction to British Association Mathematical Tables, Vol. X (Bessel Functions, Part II: Functions of Positive Integral Order). It contains, for the functions $J_n(x)$, $Y_n(x)$, $I_n(x)$ and $K_n(x)$ there tabulated, a list of formulae which includes ascending series, asymptotic expansions for functions and zeros, recurrence relations, addition formulae, derivatives, integrals, generating functions, auxiliary functions for both large and small x, and a useful summary of notations. Emphasis is on real arguments and integral orders, but some of the formulae have more general application. All physicists will welcome this compact presentation of their frequent requirements at a reasonable price, and will hope for future extensions, in particular to orders $n + \frac{1}{2}$ and arguments like $x\sqrt{i}$.

L. FOX

Notes and comments

Letter symbols, signs and abbreviations

A complex problem faced by authors and editors is the choice of letter symbols, signs and abbreviations for the various physical and chemical quantities to be used in print. They must be reasonably consistent, widely accepted in different branches of science and engineering, both in this and other countries, and moreover they must so far as possible conform to the requirements of modern type setting by machine.

A large and representative committee of publishing societies and institutions, printers and editors, has been working on this problem under the auspices of the British Standards Institution for the past two years, and the first fruits of its labours have just appeared as B.S. 1991: Part 1: 1954 *Letter Symbols, Signs and Abbreviations. Part 1 General* (Price 6s.). The many who will use this important reference book in the coming years, owe much to the Chairman of the committee, Dr. H. J. T. Ellingham, the Secretary of the Royal Institute of Chemistry, who guided and encouraged the committee through its long and difficult sessions when so many conflicting views, often strongly held, had to be resolved.

In recommending this booklet to our authors we quote from the preface, "It is more important that the general principles put forward should be accepted and followed than that there should be rigid adherence in all instances to recommendations on the use of individual symbols. Certain variations in detail may indeed be necessary in specialized fields, but there should be no departure from the general principles laid down."

We look forward to the publication of further parts of this standard which are to deal with special branches of physics and other subjects.

Napier Shaw Essay Prize

The Royal Meteorological Society announces the first competition for its Napier Shaw Prize of £100 for an original essay on "The energetics of the general circulation." The essay, which must be in English, may include material which the candidate has already published, and it may be submitted jointly in the names of more than one author. Entries must be received at the Society's Offices in London before 4 March, 1956. Fuller details may be obtained from the Assistant Secretary, Royal Meteorological Society, 49 Cromwell Road, London, S.W.7.

Optica Acta

We have received the first issue of the new international journal entitled *Optica Acta*. The English Editor is Dr. C. G. Wynne, 52 London Lane, Bromley, Kent.

The preferred languages are English, French and German, and abstracts in these three languages are given for each paper. It is intended that the journal should include papers on optical theory, optics of the eye and of vision, photometry, colorimetry, optics of instruments and other information of interest to optical workers.

In the first issue there are the following papers:

"Application des méthodes matricielles au calcul d'ensembles complexes de couches minces alternées," by Ch. Dufour and A. Herpin; "Uber die Lichtverteilung im Bild eienes Linienstückchens bei Aberrations funtionen 2. Grades in den Pupillenvariablen," by R. Krieger; "Zur Beurteilung der Bildgüte nach Definitionshelligkeit oder λ/4 Kriterium," by H. Slevogt; "The second-order aberrations in the inferferometric measurement of concentration gradients," by H. Svensson; "Une méthode permettant d'améliorer le pouvoir séparateur des instruments d'optique," by P. Lacomme; "Paraxial and third-order data of corrected doublets," by A. C. S. Van Heel; "Etude et application d'un interféromètre à polarisation," by M. Françon; "Saturation of the rod mechanism of the retina at high levels of stimulation," by M. Aguilar and W. S. Stiles; "Colloque sur les problèmes optiques de la vision," by L. Plaza.

The separate issues of the journal will not be sold singly, but in volumes consisting of 4 to 6 issues for which the price will be 3500 francs or the equivalent in other currencies. The publishers are la Société de la Revue d'Optique, 3 boulevard Pasteur, Paris, 15e.

Third general assembly and international congress of crystallography

The International Union of Crystallography has announced that the third general assembly and international congress will be held in Paris on 21–28 July, 1954.

Further information may be obtained from the secretary of the organizing committee, A. J. Rose, 1 rue Victor-Cousin, Paris, 5, and from Dr. W. H. Taylor, Crystallographic Laboratory, Cavendish Laboratory, Cambridge.

Journal of Scientific Instruments

Contents of the May issue

ORIGINAL CONTRIBUTIONS

Papers

A comparator for materials of low magnetic susceptibility. By D. A. Abbott and R. Kilgour.
New recording rain-gauges. By J. F. Nagel.
A simple method of obtaining low-temperature X-ray diffraction photographs. By S. C. Wallwork and T. T. Harding.
A microscope stage and integrating point counter for micrometric analysis of rocks. By I. H. Ford.
An analogue computing circuit for the evaluation of the ratio of two slowly-varying potentials. By R. L. Gordon.
A servo-drive for accurate speed control. By D. E. Caro, J. Y. Freeman and L. U. Hibbard.
A simple form of flat-field microscope objective. By W. S. S. Blaschke.
The construction of a combined sources unit for emission spectrography. By R. G. Stone and H. L. Bolton.
A Linde helium liquefier-cryostat combination suitable for operation down to 1·3° K. By D. H. Parkinson.
Continuous-flow cell for absorption measurements on solutions which fade. By G. P. Cooke.
A simple electromagnetic microviscometer for semi-transparent liquids. By J. Hart.
A new radiometer. By J. T. Houghton and A. W. Brewer.
Boron-loaded photographic plates as detectors of slow neutrons. By A. R. Baker.

Laboratory and workshop notes

Diffraction grating copies for electron microscope calibration. By B. E. Williams.
A simple three-terminal micrometer capacitor. By M. C. McGregor.
Calibration of a stroboscopic light. By R. Lemlich.
A thermally-operated gas valve. By E. S. Shire.
Inexpensive X-ray detectors. By U. W. Arndt.
A simple pressure controller. By D. H. James and C. S. G. Phillips.
Alginate polishing laps. By L. A. Sayce.
Soldering hard-drawn wires to terminals. By W. S. Bigg.

Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2s. 6d. including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8s. each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

SPECIAL ARTICLE

Some thoughts on the future of natural science, with illustrations from the growth of X-ray diffraction work in the United States*

By RALPH W. G. WYCKOFF, Ph.D., M.D., For. Mem. R.S., Science Attaché, U.S. Embassy, London

The current replacement of private by governmental support has been giving more impetus to applied and systematic research than to that fundamental enquiry into nature which has been the traditional basis of modern science. Some aspects of the problem for the continued development of science that result from this are discussed against a background of researches on X-ray diffraction in the United States.

No man works uninfluenced by his environment. We are all products of our times and training and of the particular society in which we live. In discussing the kinds of research with X-rays we are now carrying out in the United States, therefore, I have thought it might be of general interest to give you a rather superficial account of what we are doing, set against a background of the way in which our present work has developed as a response to American interest in the subject and to the kind of support, financial or other, that American conditions have provided. Even a subject as international as ours has developed differently here and at home; and it is only from a knowledge of the background against which our neighbours work that we can acquire some measure of that intelligent toleration and understanding of them which the world so greatly needs at this time. It is not yet true that the economic demands and rewards of society fully determine the scientific research each of us carries on, but these factors do weight the major trends of scientific development and they fix, often in their early, formative stages, the subjects which will receive most of the facilities for vigorous growth. It is this, sometimes but not always subtle, intrusion of economic considerations into the pursuit of knowledge which increasingly is setting apart the research we do from that out of which modern science has grown.

We scientists still feel it necessary to emphasize how profound has been the impact of modern science on the world's culture, how the knowledge we have acquired about natural phenomena over the last couple of centuries has multiplied the world's wealth and transformed the health and well-being of so many of its inhabitants. What we do not always bear in mind vividly enough is the profound way in which these changes are themselves altering the point of view of those of us who are providing the knowledge upon which the changes are based. There have always been people who were interested in knowledge for its own sake and others interested in knowledge they could use. The architects of modern science, and until very recently those engaged in what we are accustomed to call fundamental research, have studied nature primarily for light on the structure of the world in which they found themselves and on the true place of man in this world. The attitude of such a man towards his work is very different indeed from that of the man who seeks new knowledge for its material usefulness, whether to himself or to the general welfare. There are important national differences in the amount of sympathy and support given to those who work with the less worldly of these attitudes. Inevitably they will be less valued by the more materialistic societies. In the past they have been mainly responsible for the great advances in knowledge which more practical men have put to such good use later. The immediate future of mankind is going to be profoundly influenced by whether or not the pursuit of knowledge for the sake of understanding can continue as the materialistic goals of society are further emphasized. If it cannot, that investigation of nature which all of us have considered worth the best efforts of our individual lives will lose the services of some of man's most creative forces and may quickly assume a very pedestrian role in a developing deeper life of mankind.

It is hard for you, or any other western Europeans where a cultural tradition has become firmly entrenched over centuries, to appreciate without intimate personal experience the very different social forces that operate in a comparatively new country such as ours. When you change, as we are all so rapidly doing now, your tradition is a firm anchor which can prevent you from throwing away many of the values of your past in a momentary enthusiasm for the new thing. No new country has had time to build a firm body of tradition, and that with which we started has in some ways been weakened by the multiple origins of our present population. On top of this we have passed, over the very short period of a couple of lifetimes, from a pioneering society to one that has become wealthy and highly industrialized by applying knowledge that has already existed to the exploitation of our great reserves of natural resources. It is an added phenomenon of our time that wealth in the United States is no longer being permanently accumulated into a few hands. Our wealth is now fairly broadly distributed and has already given the "average citizen" a hitherto unattained condition of well-being. The political power is actually in his hands and since some wealth breeds in most people the desire for more one must expect from us, and from other new countries following a similar pattern of development, a continuing concern with the material side of life. This is the first time that this "average man" has had what our forbears would have considered many of the attributes of wealth, and it is inevitably exciting to him. But it brings serious problems to persons wishing to preserve and foster the cultural values which have grown up in a less equalitarian society. The pursuit of knowledge for its own sake is assuredly one of these.

Modern science arose in Europe as an effort to understand that nature which is man's environment. It has been able to flourish mainly because until recently there have been interested individuals to finance it and because the social rewards have been such as to draw good minds to its service. It is unrealistic to expect that any young society, preoccupied first with establishing itself in a hostile physical environment and then responding to the lure of the wealth easily to be gained from the natural resources found there, will place the same value on knowledge for the sake of understanding. In the United States we have had many individuals carrying out the most fundamental research, mainly in educational

* This article is a condensation of an evening discourse given before the X-Ray Analysis Group of The Institute of Physics, in London, on 20 November, 1953.

and research institutions supported by our sources of inherited wealth. But with us, as with you, such sources have rapidly been disappearing and continued support must be found elsewhere in this more democratic society. We realize that an essential ingredient of our new wealth is new natural knowledge and we are prepared to finance the search for it. But to the degree that a society really wants useful knowledge it will tend to count as successful and worthy of help that which can be promptly applied.

FUNDAMENTAL AND APPLIED RESEARCH

We have all been accustomed to distinguish between fundamental and applied research, meaning by fundamental research that which was carried out for the sake of understanding, and by applied research that which put knowledge in a useful form. The results of true fundamental research were, and are, unpredictable, whereas most applied research is, or should be, directed towards a predetermined end. But fundamental research itself has never been a homogeneous activity and as knowledge increases the disparity between its two parts becomes more marked. In the early days practically all research was exploratory. Valid results were largely unforeseen and each discovery opened new vistas into the structure of nature. Truly exploratory research still does this. It has been the work of creative individuals intent on peering into the shadow-land of the unknown and able to think and work for long periods of time in the intellectually insecure environment of the inadequately known. But each success of exploratory research opens up a large area of the now-knowable but not yet known. These areas must be surveyed if full profit is to be gained from the initial discovery. This survey of the knowable, however, interests and is the proper activity of a very different kind of man from the one who took the first step. It is a systematic affair which, having a definable object, can often be most efficiently carried out by groups of workers, rather than by individuals acting alone. There have always been these two types of basic research; what distinguishes the present situation is the overwhelming volume of the systematic research that awaits doing and the over-emphasis naturally given to it by practical providers of support who are strongly drawn to the evidently useful. Now that science has developed to the point where at any moment all imaginable experiments cannot be carried out, there is a growing danger in a society like ours that the imaginative individual doing essential exploration may not find the requisite conditions for making his unique contribution to knowledge. In this there is an obvious ultimate danger to our entire inquiry into outer nature. It is as great for the practical man as for the scholar because for both the most rewarding results have often been the least expected. By over-emphasizing the immediately useful we will ultimately dry up those sources which have most to give even to a materialistically oriented society. It is nowadays a common argument with us that the discovery of scientific fact is independent of the object for which it is sought. This ignores that what we perceive and appreciate out of the complexity of nature depends on the questions we are asking; and these are very different when we are seeking an understanding of nature from those we ask when we are seeking advances in material welfare.

CRYSTAL STRUCTURE WORK IN THE UNITED STATES

I want a little later to say something more about our current problems of supporting fundamental research, but first I should give some idea of the way crystal structure has developed in the United States, of the kinds of problems that have attracted us in the past and of our reactions to the newer problems our subject presents. The second world war brought about such profound changes in the status of our subject at home that I feel impelled to talk about first its pre-1940 history and then about crystal analysis as it exists today. In doing this I do not intend to burden you with many details of our "local history", but it is easiest to talk about the earlier stages of our work in terms of the people concerned. The subject of crystal structure determination was German and British in its origins. To me it has seemed that the German contribution remained important through the first half of the 1930's, but that, throughout, the pace of growth was in large measure set by the work of Sir Lawrence Bragg and his father, and of the many of you in this country who have been their students.

For the first 25 years following the discovery of X-ray diffraction, there were very few in the United States carrying out research in the subject. The principal reason for this was that, though towards the end of this period several universities supported it on a small scale, only the California Institute of Technology was prepared to give it at any time the larger measure of support needed to create an influential centre of activity. Experimental work was begun in three different places in the United States not so long after the publication of Sir Lawrence Bragg's first analyses of structure. In one of these, the General Electric Company's Laboratories, Hull, independently of Debye, discovered powder diffraction. Here Davey, with Hull, applied the discovery to the investigation of a number of metallic elements and simple compounds. Another early study of structure was the investigation at the Massachusetts Institute of Technology of chalcopyrite by Burdick and Ellis. The third was my own work as a graduate student at Cornell, carried out with help from Professor Nishikawa of Tokyo who happened to be visiting there. The work at the Massachusetts Institute of Technology was sponsored by Professor A. A. Noyes who left soon after to start, with Professor Millikan and Dr. Hale, the California Institute of Technology. He took Ellis and Dickinson with him and there they began work on problems of crystal analysis. Dickinson came to work with me for a brief time at the Geophysical Laboratory and I spent the following year at Pasadena where we each turned out a number of structures and thus initiated the work which has continued ever since. Pauling studied with Dickinson and assumed charge when Dickinson later turned to more strictly chemical investigations. We all know with what success he has directed it since.

None of the original group of American workers was a crystallographer and none of us had, like so many of our pre-first-war elder brothers, training abroad. Several of the first graduate students at the California Institute did thesis work in crystal analysis but as far as I know did not have opportunities to continue with it as a career. Of the next younger group of workers to enter the X-ray field in the United States, Patterson, Warren and Zachariasen all worked under the Braggs; Patterson and Pauling are the only ones among our older generation who studied in Germany. Except for Fankuchen, who also was trained here, nearly all our present mature X-ray workers are students of one or another of this older generation, or of Buerger who, around 1930, introduced X-ray methods in the Department of Mineralogy at the Massachusetts Institute.

This older group entered crystal analysis from a number of different scientific disciplines and our subject in America suffered severely from not finding general acceptance as part

of one branch of university science. With us, crystallography has never been more than an occasional course given to students of mineralogy and our mineralogists were singularly slow in realizing the value of X-ray methods. Among our older generation only Davey at Pennsylvania State College, Warren at the Massachusetts Institute, Zachariasen at Chicago and, more recently, Patterson at Bryn Mawr found places in departments of physics. Most of the rest of us were chemists by training and this in part explains why during this pre-war period Americans as a national group contributed relatively little to the methodology of our subject and were particularly interested in the ways in which a knowledge of crystal structure could throw light on rather general chemical problems. In the first days of crystal analysis we were all so astonished at the unexpected possibility of discovering where atoms were that we found adequate satisfaction in doing this for its own sake. But this stage of initial excitement soon passed and fundamental interests asserted themselves. We who were chemists began wondering how a knowledge of the atomic positions in a crystal bore on chemistry's central problem: the nature of the bond between atoms. We have continued this general interest but we have never had the sort of concentrated concern with certain groups of compounds, the metals and the silicates for instance, which has marked the progress of your research; nor have we of late made the kind of accurate determinations of bond dimensions in organic crystals which characterizes your studies at Glasgow and at Leeds.

I was from the first deeply interested in the question of how much certainty attached to the structures we could choose on the basis of the few diffraction data then obtainable; and this was intensified by a knowledge of the theory of space groups. The beautiful way the results of this purely geometrical exercise co-ordinated with the observed data brought some of us not only to use these results in deducing structures, but to take it for granted that all the atoms of a crystal were exactly in the positions thus defined by symmetry. For those of us who were unwary enough to let such an unwarranted notion take root, the evidence for "rotating" carbon chains provided by our studies of aliphatic substituted ammonium derivatives came as a serious shock. These results were indeed the beginnings of our concern with real crystals considered as less than perfectly ordered atomic arrangements. Such a realization is now, of course, an integral part of our thought; we now take it for granted that we have to ask from the outset whether or not all the atoms in a crystal are in positions determined by symmetry. Beyond that, we in America are becoming, like everybody else, increasingly involved with the complexities of order one sees so often as superlattices in alloy systems and in the kinds of defective order that express themselves through holes, Burgers vectorial displacements, and the like. We are slowly but surely moving towards a physics of the solid state which must start out from the knowledge of atomic positions that only crystal analysis can give, but which will require equally the investigation of characteristic departures from ideally perfect order before we can understand why actual solids have the properties they exhibit. Unfortunately, there is still rather a gulf between the crystal analysts in the United States and those solid state physicists who are preoccupied with crystalline defects. This cannot last indefinitely.

In considering the X-ray diffraction work we are now doing it must be remembered that though we went into the last war with relatively few workers in this field, we emerged with an activity multiplied many times. We now have a crystallographic society, the American Crystallographic Association, which has more than 700 members; most of them are concerned in some working capacity with X-rays and crystals. There are several reasons for this great expansion in research activity. Partly it was due to the demonstration, during the war, of how useful X-ray diffraction and to a lesser degree electron diffraction can be in the solution of all sorts of practical problems. Partly also it was due to the existence by this time of a backlog of well-trained young people capable of taking the jobs that were opened up by this demonstration of practical usefulness. And to this must be added the fact that American universities will rather readily add new departments when it becomes clear that the students they produce are in demand and can find jobs. But there are additional reasons. One is the recent very great increase in the amount of government support for physical research, coming largely through our armed forces. A reason of a very different sort is the widespread adoption of particle-counter techniques which can greatly abbreviate the time required for necessary observations and thus make practical all sorts of new applications, especially in industry. This introduces still another important factor that has contributed to our sudden widespread use of X-ray methods; it is the final commercial production of first rate diffraction equipment which could be used both for research and in direct industrial applications. It is not always realized how essential this is to the general adoption of a physical technique. In the period before the last war both academic and applied X-ray research in the United States was greatly hampered by the failure of our manufacturers to provide either the special tubes needed for diffraction analysis, or up-to-date versions of cameras and accessory apparatus. During the war, new, smaller companies entered the instrument-making field and the excellent X-ray tubes and equipment they have developed have been widely purchased by both universities and industrial laboratories.

I scarcely know how to give you in a few words a picture of the expanded field of research we now have after the war. Most of the people who operate our X-ray equipment, who belong to our society and who attend its meetings are interested in the practical applications of X-ray diffraction rather than in research in crystal analysis. But even without them the number of centres of academic research is greatly enlarged. Most of these are in charge of relatively young people. Important among the new centres is the Polytechnic Institute of Brooklyn where we now have, besides Fankuchen's original laboratory, the department of Ewald and the protein project of Harker. Elsewhere in the east active work is now going on in several newly organized laboratories such as Patterson's laboratory at the Cancer Institute in Philadelphia, Evans' laboratory at the Geological Survey, Mrs. Woods' laboratory at the Bell Telephone Laboratories, Parrish's laboratory at North American Philips, Hoard's laboratory at Cornell University and the newly established laboratory in the Insulation Research Laboratory at the Massachusetts Institute. The most novel of the new installations are Harker's just mentioned, and Pepinsky's at Pennsylvania State College. I am sure most of you are familiar with the purpose and make-up of both. Compared with you in this country Harker has entered late into trying to unravel the crystalline fine structure of a protein; but he has built some first rate apparatus for the purpose and he is the only one among us who has set himself up to essay a serious attack. We all hope he will have a little of that element of luck which is so essential a component of all successful research. Pepinsky's contribution to large scale research in crystal structure has been, as you well know, the development of a

remarkably complicated and well-working computational aid. It has been a very bold effort which has required a great deal of money. Many people from Europe have paid visits and worked with X.R.A.C., but thus far only a limited number of complete structures have been found with it. One will inevitably watch its productivity over the next few years because output will decide whether it represents the most effective way to handle the massive calculations that most present and future structure determinations demand. Its advantages for the earlier stages of a determination are obvious; equally obvious are the problems of making such a machine able to replace numerical computers for the final stages of an accurate determination. What is uncertain is whether it is better to have one central machine accessible to many for carrying out the early stages of their analyses or to have a number of simpler aids conveniently at hand for separate workers. Whatever the answer Pepinsky has built a fascinating machine that is a triumph of construction; and it inevitably draws attention to a far broader problem that is of primary importance to us at the moment. It is this: given adequate money for research, how much should be spent on equipment and how much for men?

It would not be feasible here to try even to list all the laboratories throughout the rest of the United States where X-ray work is going on. For a long time there has been, within the several institutions of research Pittsburg provides, an especially active interest in the study of metals. Very recently Jeffrey, coming from Leeds, has started structure work at the University there. Zachariasen's long-standing work at the University of Chicago has already been referred to; you are also familiar with Barrett's studies of metals now going on at its Institute of Metals. Seitz is at the University of Illinois and it is there too that for a quarter of a century Clark has done so much applied X-ray work. Lipscomb at the University of Minnesota and Rundle at the University of Iowa have built vigorous new schools. At the California Institute of Technology there are, of course, besides Pauling's own work, the studies of Corey and of Hughes on the amino acids and peptides. Elsewhere on the west coast X-ray departments are now to be found at the University of California, both at Los Angeles and Berkeley, and at the University of Washington.

One can scarcely give even a superficial picture of the X-ray and related work in the United States without referring to some of the fields peripheral to crystal analysis. The most obvious and long-standing of these is electron diffraction. Davisson and Germer at the Bell Telephone Laboratories were, along with Sir George Thomson, the discoverers of electron diffraction, and for some years Germer continued to study the electron diffraction from thin films. Some of our industrial laboratories, like Westinghouse and the automobile companies, which are concerned with metal surfaces have been employing it; its uses may spread, especially for non-ferrous alloys. But aside from occasional studies here and there, Brockway at Michigan maintains the only active school of electron diffraction I know about and he is no longer, as initially, concerned mainly with the investigation of the dimensions of vapour molecules. Electron diffraction equipment is now commercially available in the United States and this may increase its industrial applications. It surely has further uses, for instance in studying the rates of certain reactions and solid transformations, but one should not perhaps expect more than a moderate expansion in our use of this method.

Neutron diffraction is especially important for its potentialities in structure analysis. We are acutely aware how much it has to offer, particularly in the placing of light atoms, and we must all regret that the character of the sources it requires permits only a few to participate actively in this type of research. As you know, American contributions came originally from Shull and his associates at Oak Ridge; work now is going on at Brookhaven too.

Low angle scattering of X-rays from ordinary and macromolecular solids has been carried out in a number of places but dominant among these is the work of Fankuchen, of Bear on collagen in the Biology Department at the Massachusetts Institute and of Kaesberg at the University of Wisconsin. In the past we have had, largely through the work of Warren and his students, studies of X-ray diffraction from carbon, glasses and other non-crystalline solids. There is now less activity in this direction.

It happens that the interests of our crystal analysis and electron microscope groups overlap in their concern with electron diffraction and, therefore, one must turn also to the electron microscopists to learn all that is being done in the United States. They too are primarily concerned with X-ray microscopy. Our interest in this was started by Kirkpatrick's efforts at Stanford to make a microscope using suitably bent glass surfaces as total reflectors. This approach is technically very difficult but he has made a workable microscope that gives, I understand, a hundred or so times magnification, and development continues. When Harker was at the General Electric Company's Laboratories, he initiated experiments in extension of Kirkpatrick's; these are being continued though along other lines. We have not progressed, as has Cosslett at Cambridge, in the alternative approach of small-spot shadow microscopy. The limited interest we have shown in producing minute sources of X-rays has instead been aimed towards their use in micro-diffraction.

One of the most striking current developments in X-ray research in the United States is the rapidly expanding use of fluorescent radiation for quantitative as well as qualitative analysis. This is, of course, not at all new in principle, but it has not been much used because X-rays characteristic of most kinds of atoms in a sample could not be produced in large enough amounts for quantitative measurement after their crystalline reflexion. Even with modern tubes the method would scarcely be feasible without the sensitivity and accuracy of detection given by counter techniques.

THE RELATION BETWEEN RESEARCH AND SOCIETY

I hope this very superficial sketch of current American X-ray diffraction and allied research will give some idea of its almost explosive expansion since the beginning of the last war. I hope also that what I have said will convey to you some of the background of this expansion and of the way it has influenced both fundamental research and the range of its practical applications. Before concluding I should like to return to a few of the general relations between research and the society that supports it with which this talk began and to point out certain rather obvious interconnexions between these relations and the current growth of our subject in the United States.

As already indicated, the recent expansion of our X-ray research is not a simple and direct response to a widespread realization of how much valuable new knowledge it can produce. Rather it is the consequence of a suddenly greater financial support resulting partly from an appreciation of its practical value and partly from newly available federal money. One's initial reaction to this probably is that it is a clear indication of how urgently needed has been such a broader support for research. In a large measure this is

true but it is scarcely the whole story. Whether one wants to believe it or not the reasons why financial support is given, the mechanism whereby it is administered and the training and outlook of the personnel employed in the expanded activity it produces will all profoundly influence the future. And since many branches of physical research are now experiencing a similar expansion, it is the future of science as a whole that is being influenced. This clearly is not the place to discuss the problems thus raised, but I do want to say something about the obvious way our changed financing is influencing the balance between the two kinds of fundamental research which I earlier called exploratory and systematic.

Most of us would undoubtedly agree to the proposition that natural science can continue to flourish only if it is steadily being fed by further exploratory incursions into the unknown. What I have to say deals with the problem of how such exploration, as well as the systematic exploitation of the useful, is to be encouraged in a democratic, materialistically oriented society; and it should be emphasized that this is not a problem that is peculiarly ours. It is bound to appear wherever in the present world adequate natural and human resources create a society able to contribute to the world's culture. You in Europe, too, will have to continue facing it in order to remain in the forefront of human development. We most sincerely hope that in doing this you will arrive at a constructive solution which will not only give you a rising material standard of living but will also preserve and further that cultural outlook which has produced our western society, including modern science. Our experience already demonstrates that it is not hard to find support for systematic research. The problem of feeding science at its roots is not going to be so easy. It appears to be two-fold. On the one hand it is the problem of how to encourage individual initiative, in this case initiative of thought, in a world that by its mounting uniformity fosters accommodation to the mass. At the same time it involves persuading a democratic society to value and support an activity which, whatever its ultimate effect, certainly should not be oriented towards making an immediate and obvious contribution to that society's inevitable preoccupation with whatever enhances its material well-being. The development of crystal analysis during the post-war years illustrates some aspects of the current American reaction to this essential problem.

We all know that in the past most of the fundamental research which underlies modern science was carried out with the support of rich individuals. The only conspicuous exception to this was late 19th and early 20th century Germany where the university centres of research were state-controlled and where the Kaiser Wilhelm Gesellschaft with its outstanding Institutes was so largely financed by an enlightened industry. But this government was the antithesis of democratic, and so was the management of German industry. What it accomplished has much to tell about how to foster exploration into nature but little to indicate how democracies can meet the problem. What I have outlined as the history of X-ray research over the last few years suggests that practically all the large amount of newly-available United States Governmental money has been used for systematic, not exploratory, research. There is much talk nowadays at home about the "Project Research" in all fields that this governmental support has brought into being, about how it is swelling the scientific departments of our universities, transforming the goals of their research and putting out of balance the traditional distributions of power within these universities—while at the same time giving us a variety of magnificently equipped new governmental laboratories filled with enthusiastic but not always mature research workers. The remarks are not without some warrant. Nevertheless, those who are responsible for spending the large amounts of governmental money currently being voted for research are not finding it too hard to establish administrative systems that will give reasonably objective evaluations of systematic "Project Research" in terms of its practicability and possible contribution (near or remote) to social welfare. They have not yet fully faced the problem at the heart of the support of fundamental exploratory research. This, of course, involves choosing for support, not worthy problems for study, but those few people whose training and history of past accomplishment in exploration make them good risks for the future. Such a choice must somehow be made by one's peers but not one's competitors for support; and the selection of such peers presents difficulties that become more severe the more democratic the society. If I may be permitted to comment as an outsider I would say that one very important reason why you do so well in truly basic research is that you have inherited a workable system of selection. For the benefit of us all I, for one, hope you do not weaken it.

An inspection of the mass of new work we are carrying out in crystal analysis shows that nearly all of it is either systematic or can be made to appear so: for instance, the object of a determination of structure is clear-cut and if a reasonable choice has been made there is no undue risk of not achieving it. It is also noteworthy that the more exploratory research in our field, which centres around the individual and has broader and less sharply definable objectives, tends to have other support than the recently available federal money. Chief among such supporters are the still active research foundations and institutes established by private wealth. Sometimes it is stated that the day of usefulness of such independent sources of research money is over but I think our experience demonstrates quite the contrary. We need them more urgently than ever because they offer the only device we now have which could give exploratory research the assurance of continuity without which it cannot adequately develop. At the same time, in doing this they could be a means of discovering how this task can gradually be transferred to a society educated to an understanding of its deep importance. We must hope that these private foundations will come to realize the unique opportunity that fate is thus offering them and will concentrate their resources on carrying it out. The danger is that they will fritter away their far from unlimited funds in doing piecemeal the simpler and less hazardous jobs that government is already prepared to finance. This is especially true in the medical field where, in the past, these foundations have had such marked success. Fifty years ago there was urgent need to demonstrate that laboratory research as contrasted with clinical observation could produce a better understanding and control of human illness. This has become so apparent that the amount of public money available for the support of such medical research is sufficient for the needs of the truly limited number of gifted and trained workers our society can now supply. It is vital for the future that those who administer our remaining reserves of private endowment realize how much depends on their giving an imaginative new support to the roots of all science within the framework of a democratic society.

I am afraid many of you will be disappointed that I have not dealt in much greater detail with the X-ray work being done in the United States and for this I must apologize. But I do deeply feel that the present is especially critical in

its bearing on the future development of all natural science —that as things are now, everywhere in the world but for different reasons, science is in a real danger of being withered at its roots. This future lies partly with us as practising scientists and partly with the society in which we function. To play our part we must have clearly in mind the character of the danger; and one of the best ways to see it is in terms of that branch of science which is our central interest. I hope that what I have said will make more clear to you the present American situation, and that you will thus be led to read with sympathy what we do in the face of conditions which we are far from being alone in having to meet.

ORIGINAL CONTRIBUTIONS

The frequency response of a certain class of non-linear feedback systems

By J. C. West, Ph.D., A.M.I.E.E., and J. L. Douce, M.Sc., Electrical Engineering Department, University of Manchester

[*Paper first received* 23 *March, and in final form* 13 *October*, 1953]

A method is described for predicting the frequency response of a certain class of non-linear feedback systems of any order. The restrictions on the feedback system are: (i) only one non-linearity is present, (ii) the output of the non-linear element is a single-valued function of the input, i.e. the non-linearity is of the "simple" variety,[1] (iii) the output of the non-linear element is independent of frequency, i.e. it does not depend on the time derivatives of the input, (iv) the non-linearity is followed by an element of the system possessing a frequency-dependent transfer function of the low-pass type such as one or more integrators.[2, 3] Hence most electro-mechanical servo-mechanisms and many other cyclic control systems come within the scope of this analysis.

The "jump" phenomenon,[4–7] or discontinuity in the gain-frequency characteristic which may arise under certain conditions of input amplitude and degree of damping, is evaluated. Some experimental work on a particular system is described and the results compared with those predicted.

It has been found experimentally by many observers using various types of control systems that, with a sinusoidal input, the output of the control system is also very nearly sinusoidal and of the same frequency, although at least one element of the system has a marked non-linear characteristic. This is explained by the fact that[2, 3] in these systems the element following the non-linear element is of such a nature that all harmonics introduced by the non-linearity are attenuated to negligible amplitude at the output of the system. For instance, in a position-control system the servo motor acts as a low-pass filter, since the relationship between torque P and output angular position θ_0 takes the form

$$\begin{aligned} P &= J\ddot{\theta}_0 + F_v\dot{\theta}_0 \\ &= F_v[1 + (J/F_v)p]p\theta_0 \end{aligned}$$

where J is the total moment of inertia of the rotating parts, F_v the viscous frictional torque per unit angular velocity and p is the operational form of $\partial/\partial t$.

The filtering properties can be seen if the expression is written in the inverted form,

$$\theta_0 = \frac{1}{F_v(1 + Tp)p}P \qquad (1)$$

where $T = J/F_v$ is a time constant in seconds.

Thus higher harmonics in the torque waveform are attenuated more severely than the fundamental, and the output is very nearly sinusoidal, of the same frequency as the input.

This output signal is compared with the sinusoidal input to give the control signal, which is assumed to be approximately sinusoidal. This assumption is justified by observation of the control signal waveform, and by the agreement of the practical behaviour of the system with the theoretical response deduced from this assumption.

In linear systems the principle of superposition holds and the steady-state frequency-response characteristic is independent of the amplitude of the input signal. In a non-linear system the gain and phase-change are functions of input amplitude as well as frequency.

Again, under linear conditions the system is either stable or unstable. (In the field of control systems a system which continuously oscillates with constant amplitude and zero input is included in the class of unstable systems.) For non-linear systems it is possible for the system to be stable under certain conditions of input amplitude and frequency, including zero input, and to be unstable[8] for other conditions. In the linear system only one output is possible for a given input: in the non-linear case, however, it is possible for two or more forms of the output to exist, and for the output to jump from one to the other.

When obtaining the gain-frequency characteristic for a fixed input amplitude it often happens that, as the frequency is varied, the value of the gain "jumps" discontinuously to another level.[4, 5, 6] These gain discontinuities, or "jump" phenomena, are not necessarily undesirable and do not seem to have any corresponding effect in the transient responses obtained experimentally.[7] However, it is desirable to be able to predict the phenomenon and to show the mechanism of the jump.

RELATIONSHIP BETWEEN THE SIGNALS IMMEDIATELY BEFORE AND AFTER THE NON-LINEAR ELEMENT

The general system to be considered is shown as a block diagram in Fig. 1. The input and output signals are θ_i and θ_0 which are compared to form an error signal $\theta_i - \theta_0$. This is fed to a frequency-dependent element (marked "amplifier") which is assumed to be linear. In the steady state with a sinusoidal input, the amplifier, which may include a stabilizing network, introduces a gain $A(\omega)$ and a phase change $\phi(\omega)$. The amplifier feeds the non-linear element, the characteristic of which is assumed to be independent of frequency. For simplicity in the frequency response analysis, the signal at the input to the non-linear element will be taken as reference and of the form

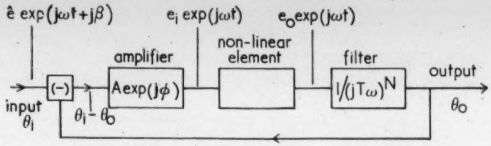


Fig. 1. Block diagram of the class of system considered

$$e_i \exp(j\omega t) \tag{2}$$

The input θ_i to the system will not in general be in phase with e_i and so the input is taken as

$$\theta_i = \hat{e} \exp[j(\omega t + \beta)] \tag{3}$$

where β is a phase angle dependent on the frequency and input amplitude. The output from the non-linear element e_0 will not be sinusoidal but only the fundamental of the waveform will be considered. In the class of non-linearity assumed for the element, the fundamental output waveform is in phase with the input signal e_i and will therefore be written

$$e_0 \exp(j\omega t) \tag{4}$$

The condition (see Appendix) for there to be no phase change of the fundamental of the output of the non-linear element is that its characteristic must be a single-valued function of the input[1]; hysteresis or backlash effects produce a phase change dependent on input amplitude.

The non-linear element is followed by a low-pass filter so that it is justifiable to neglect the higher harmonics. Thus the output of the system is given by

$$\theta_0 = [1/(j\omega T)^N] e_0 \exp(j\omega t) \tag{5}$$

There are two relationships that can be found between e_0 and e_i, both of which must hold at the same time. The first is formed from the non-linear characteristic and is the amplitude of the fundamental component of the output waveform as a function of input amplitude. This relationship is independent of frequency and the results can be plotted as a graph of e_i against e_0.[3]

The second relationship takes account of the loop gain of the rest of the system which is assumed to be completely linear. Thus the output e_0 from the non-linear element passes through the filter and is combined with θ_i to form the error signal which passes through the amplifier to form e_i. Thus there is a relationship between e_i and e_0 which does not involve the non-linear characteristic, but depends on the transfer functions of linear elements and also on the frequency and amplitude of the input signal. It can be obtained experimentally by removing the non-linear element, thus opening the loop, and injecting a signal e_0 of the same frequency as the input θ_i and adjusted to be the same phase as e_i. When the loop is closed both relationships must be obeyed and thus the operating conditions occur where both relationships give the same values for e_i and e_0.

If the output from the non-linear element is $e_0 \exp(j\omega t)$ and the output of the system is

$$\theta_0 = [e_0/(j\omega T)^N] \exp(j\omega t) \tag{6}$$

[See equations (4) and (5)]

$$\text{Then} \quad \theta_i - \theta_0 = \left[\hat{e} \exp(j\beta) - \frac{e_0}{(j\omega T)^N}\right] \exp(j\omega t) \tag{7}$$

The input to the non-linear element is therefore

$$e_i \exp(j\omega t) = \left\{\hat{e} \exp[j(\beta + \phi)] - \frac{e_0}{(j\omega T)^N} \exp(j\phi)\right\} A \exp(j\omega t) \tag{8}$$

This becomes

$$e_i = \left[\hat{e} \cos(\beta + \phi) - \frac{e_0}{(\omega T)^N} \cos\left(\phi - \frac{N\pi}{2}\right)\right] A \tag{9}$$

because the imaginary part of the right-hand side of equation (8) must be zero, i.e.

$$\hat{e} \sin(\beta + \phi) = \frac{e_0}{(\omega T)^N} \sin\left(\phi - \frac{N\pi}{2}\right) \tag{10}$$

From this one obtains

$$\cos(\beta + \phi) = \pm\left[1 - \left(\frac{e_0}{\hat{e}}\right)^2 \frac{1}{(T\omega)^{2N}} \sin^2\left(\phi - \frac{N\pi}{2}\right)\right]^{\frac{1}{2}} \tag{11}$$

The phase angle β can be eliminated from equations (9) and (11) giving

$$\frac{e_i}{\hat{e}} = A\left\{\pm\left[1 - \left(\frac{e_0}{\hat{e}}\right)^2 \frac{1}{(T\omega)^{2N}} \sin^2\left(\phi - \frac{N\pi}{2}\right)\right]^{\frac{1}{2}} - \frac{e_0}{\hat{e}} \frac{1}{(T\omega)^N} \cos\left(\phi - \frac{N\pi}{2}\right)\right\} \tag{12}$$

THE RESPONSE OF A SECOND-ORDER SYSTEM

The particular system to be considered in this section is a simple second-order remote-position-control servo-mechanism, incorporating saturation of the control amplifier. Stabilization is effected by derivative of error damping. Representing the transfer function of the stabilizing unit and amplifier by

$$g(1 + pT_s) \tag{13}$$

it follows that, from Fig. 1

$$A = g(1 + \omega^2 T_s^2)^{\frac{1}{2}} \tag{14}$$

$$\sin \phi = \omega T_s (1 + \omega^2 T_s^2)^{-\frac{1}{2}} \tag{15}$$

With $N = 2$ and substituting for A and ϕ from equations (14) and (15), equation (12) becomes

$$\frac{e_i}{\hat{e}} = \frac{g}{\omega^2 T^2}\left(\frac{e_0}{\hat{e}}\right) \pm g\left[(1 + \omega^2 T_s^2) - \left(\frac{e_0}{\hat{e}}\right)^2 \left(\frac{\omega T_s}{\omega^2 T^2}\right)^2\right]^{\frac{1}{2}} \tag{16}$$

This equation is plotted in Figs. 2 and 3 for several values of input frequency; Fig. 3 is for a lightly damped system, and Fig. 2 represents a system critically damped in the linear regime.

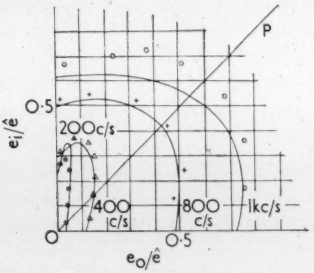


Fig. 2. The characteristics of the linear part of a heavily damped system

$g = 0{\cdot}165$; $T = 1{\cdot}35 \times 10^{-4}$ sec; $T_s = 2{\cdot}22 \times 10^{-2}$ sec.
Theoretical curves are shown by full lines.
Experimental results: ⊗ = 200 c/s; △ = 400 c/s; + = 800 c/s; ○ = 1 kc/s.

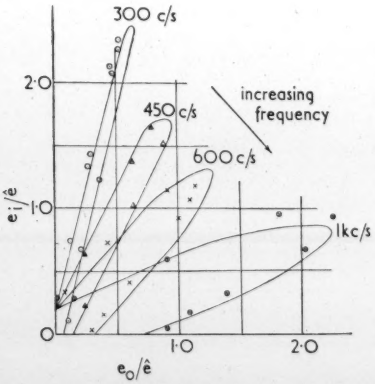


Fig. 3. The characteristics of the linear part of a lightly damped system

$g = 0{\cdot}258$; $T = 1{\cdot}35 \times 10^{-4}$ sec; $T_s = 5{\cdot}37 \times 10^{-5}$ sec.
Theoretical curves are shown by full lines.
Experimental results: ○ = 300 c/s; △ = 450 c/s; × = 600 c/s; ⊗ = 1 kc/s.

From equation (16) it is seen that if $e_0/\hat{e}$ is zero then $e_i/\hat{e}$ increases with frequency. This is clearly shown in Fig. 2 for the critically damped system; for the lightly damped system of Fig. 3 this effect is not appreciable at the frequencies considered.

The saturation characteristic is taken to be linear and of unity gain for inputs less than $\pm h$, and to saturate for larger inputs with an output level of $\pm h$.

Thus
$$e_0 = e_i \quad \text{for} \quad -h \leqslant e_i \leqslant h$$
$$e_0 = h \qquad e_i > h$$
$$e_0 = -h \qquad e_i < -h$$

For a sinusoidal input, $mh \sin wt$ where $m > 1$, the output waveform is analysed into Fourier components and the amplitude of the fundamental is found to be

$$B_1 = 2h/\pi\left\{m \sin^{-1}(1/m) + \sqrt{[1 - (1/m)^2]}\right\} \tag{17}$$

The complete fundamental of output-input relationship is thus as shown in Fig. 4. As the input becomes very large, the amplitude of the fundamental of the output waveform tends towards $4h/\pi$ which is the amplitude of the fundamental of a square wave of peak amplitude h.

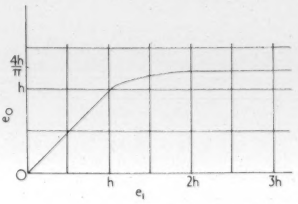


Fig. 4. The describing function for the non-linear function considered

The required relationship between $e_i/\hat{e}$ and $e_0/\hat{e}$ is shown in Fig. 5, for three values of input amplitude, δ being $\hat{e}/h$. For a small input of peak magnitude h, saturation begins for $e_i/\hat{e} = e_0/\hat{e} = 1$ and the maximum value of $e_0/\hat{e}$ is $4/\pi$, this condition corresponding to curve (i) in Fig. 5. As the input amplitude θ_i is increased, the maximum value of $e_0/\hat{e}$ decreases;

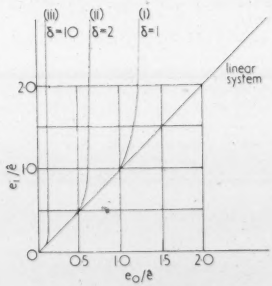


Fig. 5. The variation of the describing function with input amplitude (δ)

curve (ii) corresponds to $\delta = 2$ and curve (iii) applies for $\delta = 10$.

Evaluation of the response. The steady-state operating conditions of the system are determined by the intersection of the curves of Fig. 5 with those of Fig. 2 or Fig. 3. Having found the equilibrium value of $e_0/\hat{e}$, the gain of the system, $\left|\frac{\theta_0}{\theta_i}\right|$ is, by equation (6)

$$\left|\frac{\theta_0}{\theta_i}\right| = \left(\frac{1}{\omega T}\right)^2 \left(\frac{e_0}{\hat{e}}\right) \tag{18}$$

Considering Fig. 2 for the heavily-damped system, it is found that, at any particular values of input amplitude and frequency, the two curves intersect at one point only, so that there is no ambiguity concerning the operating values of e_i and e_0.

For the lightly damped system, Fig. 3, it is seen that, over a limited range of input amplitudes and frequencies, there are three points of intersection, so that further information is required before the operating point can be determined uniquely. This information may be deduced by consideration of Fig. 6 which shows the describing function of the non-linear element for one input amplitude, with the loop response of the linear system for several frequencies.

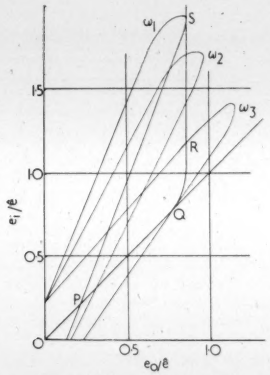
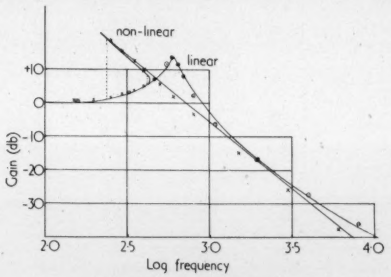


Fig. 6. Determination of the steady state operating point for fixed input amplitude

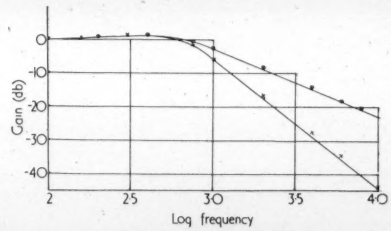
Let the frequency of the input sinusoid be gradually increased, from an initial value less than ω_1. As the frequency is increased from a low value, the point of intersection of the locus and describing function moves along the line OQ. When the input frequency is equal to ω_3, the operating conditions are represented by the point Q, and it is evident that, on further increasing the frequency of the input sinusoid, there will be a discontinuous increase in the values of $e_i/\hat{e}$ and $e_0/\hat{e}$ since the operating conditions will transfer to near the point R. As the frequency is increased still further, $e_0/\hat{e}$ approaches its maximum value. Thus, from equation (18) the gain will eventually fall with the square of the frequency. For high values of input frequency, the sole operating point is along the upper limb of the locus, and as the frequency is reduced, this point moves from R, at frequency ω_3 to S for $\omega = \omega_1$. For the next decrease in frequency there is a large discontinuous decrease in $e_i/\hat{e}$ and $e_0/\hat{e}$, and the system now operates in the linear regime (point P).

Thus with the type of non-linearity considered, it is expected that, for a lightly-damped system with an input magnitude lying within certain limits, there will be a discontinuity or "jump" when the frequency-response is being taken. On decreasing the frequency, the jump will be larger than that which occurs when the frequency is being increased, and there will be a "hysteresis" region, where two outputs are possible for a given input, the operating conditions depending on the past history. Fig. 7(*a*) and (*b*) show the theoretical



(*a*). A lightly damped system

Theoretical curves are shown by full lines.
Experimental results: ○, linear system; ×, non-linear system, $\delta = 3$.



(*b*). A heavily damped system

Theoretical curves are shown by full lines.
Experimental points: ○, linear system; ×, non-linear system, $\delta = 3$.

Fig. 7. Experimental and predicted frequency response curves

frequency-response of the system with different amounts of damping, both for linear operation and for the case where saturation is present.

The theory also predicts that, over a range of input frequencies, there will be discontinuities in the input-output graph for a constant-frequency input. This may be seen from Fig. 8 which shows the locus for one particular frequency with the describing function of the non-linear element for several values of input amplitude. As the input amplitude is increased from zero, the operating point is constant at P, so that the gain of the system is constant. This applies until $\delta > 5$ when the curves no longer intersect in the region of P,

so that the operating point moves to R. Further increase in input amplitude merely decreases the gain and the output tends to remain of constant amplitude.

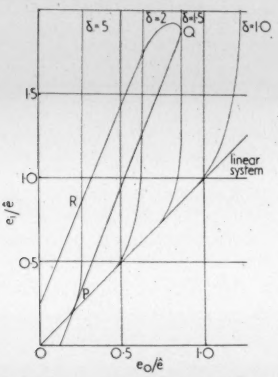


Fig. 8. Effect of varying input amplitude at a constant frequency

On decreasing the input amplitude, the operating point moves along the upper limb until there is no point of intersection of the locus and describing function in this region. This occurs for $\delta = 1{\cdot}5$ when the operating point reverts to P, and the gain is then constant. The amplitude-response of the lightly-damped system is shown in Fig. 9, for three different frequencies.

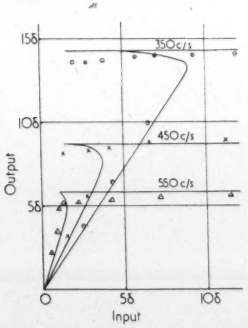


Fig. 9. Operating conditions at fixed input frequency

Theoretical curves are shown by full lines.
Experimental results: ○ = 350 c/s; × = 450 c/s; △ = 550 c/s.

EXPERIMENTAL RESULTS

To render the experimental system as flexible as possible, an electronic servo-simulator[9] has been used to verify the theoretical results, the servo motor being represented by two high-gain Miller integrators. The time scale has been adjusted so that the natural frequency is approximately 600 c/s enabling audio-frequency apparatus to be used in conjunction with the system. The saturation characteristic is introduced by two diodes biased to appropriate voltages. Fig. 10 shows the block diagram of the complete simulator. The experimental frequency-responses for the system (*a*) lightly damped,

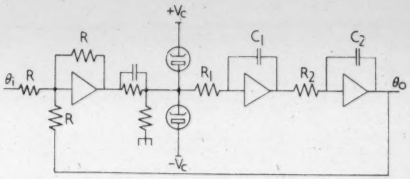


Fig. 10. The schematic circuit of the simulator employed

$T_1\omega_0 = 0{\cdot}2$, and (*b*) approximately critically damped, $T_1\omega_0 = 2{\cdot}1$ are shown in Figs. 7 (*a*) and (*b*). These agree with the theoretical curves both for the linear case and for the system with saturation present.

The experimental amplitude-response of the underdamped system is shown for comparison with the theoretical results of Fig. 9. That portion of the curves of Fig. 9 over which an increase in input gives a decrease in output is essentially unstable, and consequently the experimental result is a sharp jump in output when this region is reached.

The experimental loci for the relationship between e_i and e_0 for the open loop have been determined for several frequencies. Whilst the most fundamental method of obtaining these loci is to open the loop by removing the non-linear element, and injecting a signal of variable amplitude and of the same phase as e_i, this method presents several practical difficulties. The adjustment of the two phases is a tedious operation, since the phase of e_i changes appreciably as the phase of e_0 is adjusted. Opening the loop allows the high-gain integrators to drift in a random manner, and trouble has also been encountered due to pick-up from the supply mains. A more satisfactory method is to replace the non-linear element by a variable-gain, zero-phase-change amplifier, and to record simultaneous values of e_i and e_0 as the gain is varied. Referring to Fig. 2 the effect of varying the gain of the loop is to vary the slope of the line OP, and, as the gain is increased from zero, the slope continuously decreases from being infinite, so that the complete relationship between $e_i/\hat{e}$ and $e_0/\hat{e}$ is described for one particular frequency.

The experimental loci obtained in this manner are plotted on Figs. 2 and 3 for the system with two different degrees of damping. These results show agreement with the theoretical loci apart from the variation at very low frequencies, where the maximum value of $e_i/\hat{e}$ for the practical system tends to zero, whilst the theoretical maximum tends to infinity as the frequency is continually reduced.

This departure from theory for the low-frequency response is accounted for by the imperfect integrators used in the simulator. The integrating stages have a finite gain at zero frequency and the correct expression for the transfer function of the low-pass filter (Fig. 1) is

$$\theta_0/e_0 = [G/(1 + pT)]^2 \qquad (19)$$

where G is the zero frequency gain of one integrator, and T is the equivalent time constant. Each integrator has an

open loop gain of 800 and it may be shown that this modifies the loci for frequencies below 300 c/s.

STABILIZATION BY VELOCITY FEEDBACK

The block diagram of a second-order servo mechanism stabilized by velocity feedback is shown in Fig. 11. Using the notation indicated in the figure,

$$e_i \exp(j\omega t) = \hat{e} \exp[j(\omega t + \beta)] - \frac{e_0}{\omega^2 T^2} \exp(j\omega t) - \frac{K}{j\omega T} e_0 \exp(j\omega t)$$

giving

$$\frac{e_i}{\hat{e}} = \frac{e_0}{\hat{e}} \cdot \frac{1}{T^2\omega^2} \pm \left[1 - \left(\frac{K}{\omega T}\right)^2 \left(\frac{e_0}{\hat{e}}\right)^2\right]^{\frac{1}{2}} \qquad (21)$$

This expression is very similar to that for the system stabilized by phase-advance, the significant difference being that $e_i/\hat{e}$ is constant for all frequencies for $e_0/\hat{e}$ equal to zero.

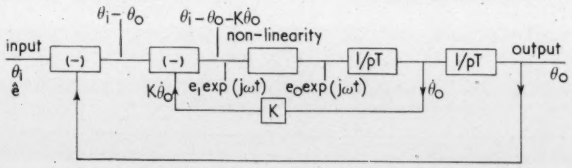


Fig. 11. Simulation of a simple remote-position-control servo-mechanism stabilized by velocity feedback

ADDITIONAL PHENOMENA

The type of characteristic considered in this paper has an incremental gain or slope which decreases as the error increases, reaching zero in the case of true saturation. This can be classed as "soft-spring" and it is seen that when the jump phenomenon occurs it is always an increase in amplitude for increasing frequency. Similar work has been carried out on "hard-spring" characteristics,[10] i.e. incremental gain increasing with error signal, and the jump phenomenon again occurs but in the opposite direction. A decrease in amplitude takes place for an increase in frequency.

Experimental work with unsymmetrical non-linear characteristics is being pursued. Here it has been found, in addition to discontinuous jumps in the output amplitude, there is also the sudden appearance of subharmonics.

CONCLUSION

A method has been described by which the frequency-response of feedback systems containing one single-valued non-linear element followed by a low-pass filter, may be ascertained. This method gives a graphical presentation of the mechanism of the "jump" phenomenon or gain discontinuity.

It is seen from Fig. 8 that a jump can only occur if the frequency-dependent locus cuts the non-linear describing function at least three times.

The agreement between experiment and theory justifies the use of the describing function technique.

REFERENCES

(1) CHERRY, E. C., and MILLAR, W. *Automatic and Manual Control*, p. 263 (London: Butterworths Scientific Publications, Ltd., 1952.)
(2) KOCHENBURGER, R. J. *Trans. Amer. Inst. Elect. Engrs*, **69**, p. 270 (1950).
(3) JOHNSON, E. C. *Trans. Amer. Inst. Elect. Engrs*, **71**, p. 169 (1952).
(4) SCOTT, W. E. *Automatic and Manual Control*, p. 249 (London: Butterworths Scientific Publications Ltd., 1952.)
(5) PIPES, L. A. *Applied Mathematics for Engineers and Physicists*, p. 603 (New York: McGraw Hill Book Co. Inc., 1946).
(6) LUDEKE, C. A. *J. Appl. Phys.*, **17**, p. 603 (1946).
(7) HAMMOND, P. J. *T.R.E. Technical Report*, No. 178 (*Dec.*, 1952).
(8) SHERRARD, E. S. *Trans. Amer. Inst. Elect. Engrs*, **71**, p. 312 (1952).
(9) WILLIAMS, F. C., and RITSON, F. J. U. *J. Instn Elect. Engrs*, **94**, Part IIa, p. 112 (1947).
(10) WEST, J. C., and NIKIFORUK, P. *Proc. Instn Elect. Engrs*, **101**, Part II (1954). To be published.

APPENDIX

Let the characteristic of the non-linear element be given by

$$e_0 = f(e_i)$$

where $f(e_i)$ is a single-valued function of the input e_i.

If the input is sinusoidal of the form

$$e_i = B \cos \omega t$$

then the output is

$$e_0 = f(B \cos \omega t) = f'(t)$$

This is a periodic waveform and can be expressed as the series

$$f'(t) = a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \ldots + b_1 \sin \omega t + b_2 \sin 2\omega t + \ldots$$

Since $f(e_i)$ is a single-valued function of e_i and $\cos \omega t$ is an even function of t then $f'(t)$ is an even function of t,

i.e.

$$f'(-t) = f'(t)$$

Replacing t by $-t$ in the series shows that $b_1, b_2 \ldots$ are all zero.

The fundamental component is $a_1 \cos \omega t$ and is in phase with the input signal. Thus a sufficient condition for zero phase shift of the fundamental is that the characteristic shall be single-valued, i.e. no hysteresis.[1] If, further, the characteristic is symmetrical about the origin so that

$$f(e_i) = -f(-e_i)$$

it can easily be shown that, in addition to the above, all the even harmonics are zero,

i.e. $a_0, a_2, a_4 \ldots$ are zero.

Spectrophotometry of fluorescent pigments

By the late R. DONALDSON, M.A.,* National Physical Laboratory, Teddington, Middlesex

[*Paper first received* 15 *September, and in final form* 12 *October*, 1953]

A method is described for measuring the spectral reflexion of fluorescent pigments for incident monochromatic light covering the ultra-violet and visible regions of the spectrum. The results are recorded in the form of a square table from which quantum efficiencies and spectral reflexion curves for sunlight or any illuminant of known energy distribution are calculated. Data are given for a red pigment and an optical bleach. Three curves only, showing the variation of tristimulus values with incident wavelength, are sufficient for the calculation of colour. The curves of tristimulus values may be derived from the reflexion curves or measured directly by a template colorimeter.

When an attempt is made to measure the spectral reflexion of fluorescent materials on the ordinary type of spectrophotometer the results are misleading. If the sample is placed between the monochromator and the photocell, the fluorescent light, excited by the monochromatic light, reaches the photocell undispersed and is erroneously recorded as reflected monochromatic light. This is what happens with the General Electric recording spectrophotometer and also with the Beckman, Uvispec and Unicam instruments. By illuminating the sample first by white light, accurate results can be obtained, but they apply only to that illuminant. None of the commercial spectrophotometers are adapted to measure specimens in this position, and there is also the difficulty of finding the appropriate illuminant, because it is the appearance in daylight which matters with fluorescent pigments used for advertising and display. The colorimetric illuminants *B* and *C* are a good substitute for daylight, as far as the visible spectrum is concerned, but they were not designed to cover the ultra-violet. An artificial daylight illuminant with the correct ultra-violet content is still lacking.

In the determination of the standard colour cards of the Textile Colour Card Association (T.C.C.A.), Reimann, Judd and Keegan[1] drew attention to the presence of appreciable fluorescence in about 25% of the colour cards, and found that straightforward measurements on the General Electric spectrophotometer were unreliable for such specimens. Later, Tyler and Callahan[2] measured some strongly fluorescent pigments with a modified General Electric spectrophotometer. A blue-green filter was placed over the photocell of the spectrophotometer to eliminate the fluorescence, and reflexion curves were taken with and without the filter. The limitations of this method are that the fluorescence is regarded as substantially monochromatic, and that the effect of the ultra-violet cannot be taken into account.

The investigation described here is concerned with modifying spectrophotometric methods to deal with fluorescent pigments. Two methods are shown. The first is a complete analysis, in which the distribution of the reflected light is found for all incident wavelengths. The second, a shortened method, gives the colour of the pigment directly.

APPARATUS

The apparatus used in this investigation is much more elaborate than that normally required for ordinary colour measurement. The arrangement is shown diagrammatically in Fig. 1. The fluorescent sample is illuminated by monochromatic light provided by a double monochromator, a Müller–Hilger instrument with quartz optics. A single instrument would serve, provided the stray light is measured by the second monochromator and corrections applied. The double monochromator was available however, and the necessity for correcting for stray light was avoided.

The light reflected from the specimen is analysed by a second monochromator. This analysing monochromator is, in fact, a template colorimeter with a template in the plane of the spectrum. When a slit, which can be moved laterally, is substituted for the template, the instrument functions as a double monochromator. One unusual feature about this colorimeter is that it employs the same prism to disperse and to recombine the light. Firstly the light is dispersed into a spectrum by a 60° prism. The spectrum is then imaged above itself by means of a lens and two mirrors at right angles. By means of a field lens near the plane of the spectrum, the prism surface is imaged at the reflecting mirrors and then back on to itself so that the prism is completely filled by the returning beam. After recombination at the exit slit the light is received by a photocell. The stray light is eliminated by placing the receiver shutter after the first, but before the second dispersion (see Walsh[3] and Donaldson[4]).

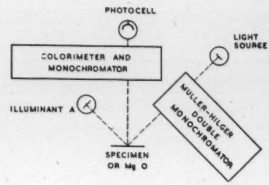


Fig. 1. Arrangement of the apparatus

The chief experimental difficulty resulting from the use of two monochromators in series is lack of light. It is therefore necessary to conserve the light as much as possible when it is being diffused by the sample, for it is there that the loss of light is greatest. The slit of the first monochromator is focused on the sample by means of a concave mirror, and the illuminated area is then imaged on the slit of the analysing monochromator by a long focus microscope objective of high aperture. Thus, the cones of light used in the illumination and in the viewing of the sample were as wide as the experimental conditions and available apparatus would allow. The optical arrangement is not shown in the diagram.

When fluorescence is present, a difficulty concerning units arises which is absent in ordinary spectrophotometry. In non-fluorescent reflexion the light reflected is of the same wavelength as the incident light so the ratio of light intensities is easily determined. Where there is fluorescence, however, one wavelength excites another and a common unit for both radiations is required. The obvious unit is relative energy. This requires that the analysing monochromator and photocell be calibrated in terms of spectral energy. For the visible spectrum a tungsten lamp of known colour temperature

* Owing to the untimely death of Mr. Donaldson the proofs of this paper were passed by Dr. W. S. Stiles.

(illuminant *A*) forms a suitable standard of energy distribution. A magnesium oxide screen, as shown in Fig. 1, is illuminated by the tungsten lamp and the response of the photocell-monochromator unit is noted throughout the spectrum. From the data on the energy distribution of illuminant *A*, factors are derived which convert the readings given by the photocell into units of relative energy. This method fails in the ultra-violet so a thermopile was used for the range 300–400 mμ. Here the illuminant was a mercury arc and the energy in the spectral lines was measured directly.

PIGMENTS

The fluorescent pigments examined were taken from the current commercial catalogues of fluorescent inks. These pigments are prepared by dispersing organic dyes in synthetic resin, the efficiency and permanence being closely dependent on the degree of dispersion. A very large number of different colours are shown in the catalogues but, as they can obviously be classed into a few groups, a representative selection was made and measured. All the inks show somewhat similar behaviour, so the red ink was chosen as it seemed to illustrate the main features most clearly and a complete set of curves is given for this ink.

To prepare the ink sample for measurement, a small circular piece of the material was cut out from a page of the catalogue, and mounted by a touch of adhesive to a flat metal disk of the same diameter. The ink sample and the magnesium oxide screen were attached to opposite ends of a rod which could be rotated about its middle point, thus allowing an interchange of magnesium oxide and ink sample.

RESULTS FOR INCIDENT MONOCHROMATIC LIGHT

Fig. 2 illustrates what happens when the red ink is illuminated by monochromatic light. The wavelength of the incident light is 500 mμ and the energy content is shown by the area of the column, 100 units high and base 200 mμ. The amount of the fluorescence, as shown by the continuous

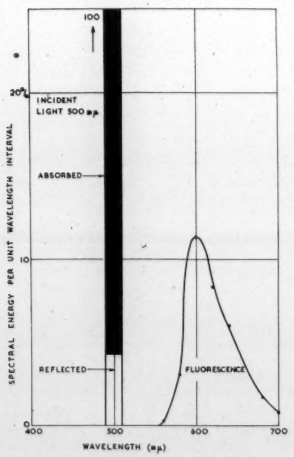


Fig. 2. Fluorescent ink (red) illuminated by monochromatic light of wavelength 500 mμ

curve, is in proportion to this. The amount of the incident light absorbed is represented by the blackened portion of the column, the remainder being reflected without change of wavelength. The term "unmodified reflexion" will be applied to the reflexion without change of wavelength to distinguish it from the fluorescent reflexion. The fluorescence curve only reaches a height of 10%, but it covers a fair wavelength range, consequently the area under the curve is a sizable fraction of the area of absorption.

In Fig. 3 is shown the result for incident light of wavelength 580 mμ. The incident light is now very close to the

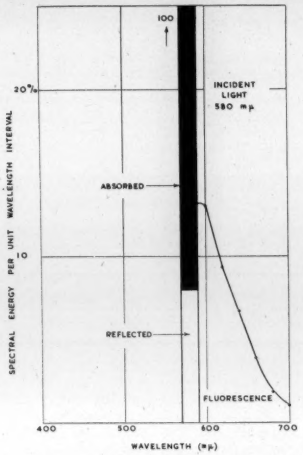


Fig. 3. Fluorescent ink (red) illuminated by monochromatic light of wavelength 580 mμ

band of fluorescence, but there is no sign of any diminution of fluorescence yield. Once the excitation wavelength is within the band, the fluorescence vanishes and only the unmodified reflexion remains. The critical wavelength at which this takes place could not be determined exactly on account of the width of the slits being rather large—each monochromator had a semi-band width of 10 mμ—but the indication was that the fluorescence fell suddenly. Fig. 3 shows the fluorescence curve as stopping abruptly, when the rectangle indicating the incident light is reached. No measurement of the fluorescence could be made below 600 mμ because the slits were too wide to allow a clear separation of the excitation and fluorescence radiations. The probability is, however, that the fluorescence band retains its shape.

CONSTANT QUANTUM EFFICIENCY

The efficiency of a pigment depends on two quantities, the unmodified reflexion and the amount of fluorescence that is added to it. The additional increase due to the fluorescence is the important factor, for high ordinary reflexions are common. A comparison of the relative amounts of fluorescence and absorbed light can be made in energy units, or more appropriately in light quanta. To convert energy into quanta requires a multiplying factor proportional to the

wavelength, the whole of the fluorescent band being integrated. Thus, if the sample absorbs 100 quanta at any wavelength the number which reappears as fluorescent light is termed quantum yield or quantum efficiency. Some of these inks have the interesting property of possessing a constant quantum efficiency for a considerable range of incident wavelengths. Fig. 4 shows the results for the red ink. From the shortest

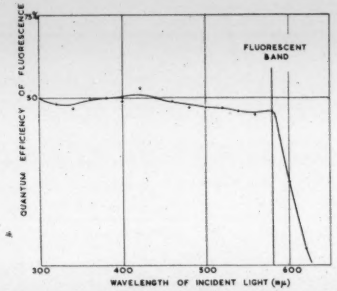


Fig. 4. Quantum efficiency of fluorescent ink (red) for absorbed light

wavelength used (300 mμ) upwards to the fluorescent band, the curve is flat and the quantum efficiency has a nearly constant value of about 50%. When the incident radiation is nominally 600 mμ, the fluorescence diminishes to about half this value. This is probably due to only half of the wide band of incident radiation being effective. The values shown in the range 300–400 mμ were not derived directly from measurements with ultra-violet radiation of corresponding wavelengths. As mentioned earlier, the mercury lines were used in this range. The ordinates of the fluorescence curves for the uniform spacing of the wavelengths were obtained by graphical interpolation. Values at equal wavelength intervals are more convenient for carrying out the integrations required later.

The phenomenon of constant quantum efficiency is well known and has been investigated for a number of substances by Wawilow,[5] Bowen,[6] Johnson, Watanabe and Tousey[7] and others. Greater accuracy than the apparatus could give would be required to decide whether the curve is exactly flat and the irregularities shown in the diagram result from experimental error. Very precise measurements of efficiency would be of some importance because fluorescence measured by photocell can be made the basic principle of a neutral receiver for evaluating radiation (see Bowen and Sawtell[8]). The receiver differs from the thermopile in that it responds to number of quanta and not to energy.

The curve of Fig. 4 refers only to the light absorbed. A small amount of the incident light is always reflected without change of wavelength. When this is added to the fluorescent light, the result is shown in Fig. 5, the 100% line now being

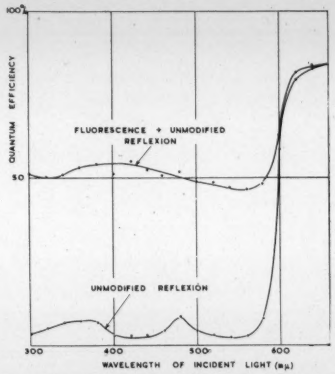


Fig. 5. Quantum efficiency of fluorescent ink (red) for incident light

the incident light. The presence of this small and rather irregular amount of unchanged reflected light destroys the ideal flatness of the quantum efficiency curve. The design of a neutral radiation receiver, operating on the fluorescence principle, is complicated by this unmodified reflexion. It would be necessary to utilize the unmodified light by reflecting it back on to the fluorescent sample or to reduce it to a negligible amount.

TABULAR DESCRIPTION

The reflexion of a non-fluorescent material is described by a single spectrophotometric curve. A fluorescent material, on the other hand, requires a complete curve for each incident wavelength. It is impossible to show clearly on a two-dimensional diagram all the spectral distributions of fluorescent light arising from incident light in the ultra-violet and the visible, but a complete record can be conveniently arranged as a square table as shown below. The wave-

Reflexion of fluorescent ink (red)

Incident light (mμ)	*Spectral reflexion (%) mμ*											
	300	340	380	420	460	500	540	580	620	660	700	740
300	3·5							1·9	5·6	2·1	0·5	0·1
340		6·7						1·6	5·8	2·2	0·6	0·1
380			6·8					2·3	7·0	2·5	0·8	0·2
420				2·8				2·9	8·6	3·2	1·0	0·2
460					4·4			2·7	8·3	3·4	0·8	0·2
500						4·3		3·0	8·3	3·4	0·8	0·3
540							2·4	3·7	8·8	3·6	0·9	0·3
580								8·0	9·3	3·9	1·0	0·4
620									79·5	0·4	0·1	—
660										84·1	—	—
700											85·1	—
740												86·0

lengths of the incident light are shown in the first vertical column, and the quantities in energy units are equal. Each row of the table indicates the amount of unmodified reflexion and fluorescent reflexion as a percentage of the incident light. The unmodified reflexion thus finds itself on the diagonal of the table. In accordance with Stokes' law the entries for fluorescence are always in the top half of the table. Only half of the total number of wavelengths is shown but, even with all the entries, the table is not unwieldy or impracticable.

From a table of this kind it is possible to derive the reflexion curve for any illuminant of prescribed energy distribution. Each row is multiplied by the appropriate energy contribution of the illuminant and the results are summed. This assumes that the fluorescence varies linearly with the intensity of the excitation radiation and that the contributions of all the wavelengths are additive. For this kind of fluorescent material the requirement seems to be satisfied for moderate light intensities. The linearity was tested over a range of ten to one and no measurable deviation from proportionality was found. It almost certainly follows that if the linearity is correct the additivity will also hold.

SPECTRAL REFLEXION FOR VARIOUS ILLUMINANTS

The spectral reflexion curve for illuminant *A* can be calculated from a table as above and it can also be measured directly, thus providing a check on the whole method. Reasonable agreement with all the samples was obtained showing that no important factor had been neglected. One source of error, however, which was at first overlooked, was revealed by this check. It was found that, when the slit is focused on the fluorescent sample, a certain amount of spreading of the image takes place and care must be taken that the slightly enlarged image is not being intercepted by the slit of the second monochromator. This can happen if the slit of the second monochromator is adjusted to admit

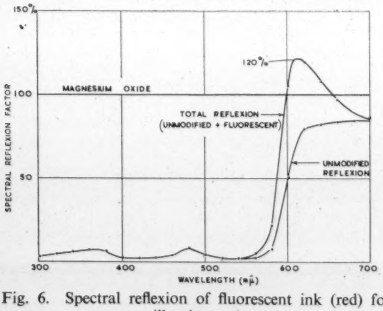


Fig. 6. Spectral reflexion of fluorescent ink (red) for illuminant *A*

narrowly the image from the magnesium oxide screen. Fig. 6 shows the calculated result for illuminant *A*. The straight line, value 100%, represents magnesium oxide. The other curves show the unmodified reflexion and the total reflexion with the fluorescence added. Thus, for incandescent light, an apparent reflexion of about 120% is reached at 610 mμ.

In Fig. 7 are shown similar results for sunlight calculated from the average energy distribution as given by Moon.[9] Here the proportion of fluorescence is very much higher than for illuminant *A*. The curve reaches a peak value of 165%. With north skylight, which has relatively more blue light, the peak value of the reflexion would be still higher. Of all the inks examined, the red showed the highest reflexion. The yellow, which looked very brilliant, only reached a value of 102% for illuminant *A*. The very high efficiency of the red

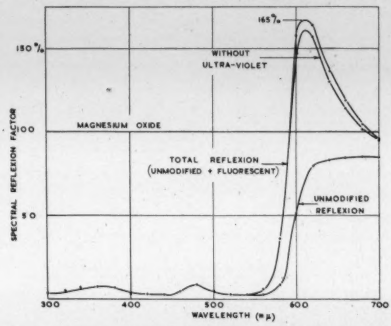


Fig. 7. Spectral reflexion of fluorescent ink (red) for sunlight

is due to two factors. Firstly, it has a high unmodified reflexion in the region where the fluorescence is, and secondly, the exciting radiation is drawn from a wider range of wavelengths than for the other colours.

The middle curve in Fig. 7 refers to sunlight with the ultra-violet removed. The whole effect of the ultra-violet accounts for only about 10% of the fluorescence. It is the visible light which is producing most of the fluorescence, not the ultra-violet as might be expected.

OPTICAL BLEACHES

Another kind of fluorescent material, widely used domestically, is the optical bleach or fluorescent brightener. It is added to soap powders and detergents to enhance the whiteness of laundered clothing. Fig. 8 shows that the fluorescence is in the blue region of the spectrum, the exciting wavelength

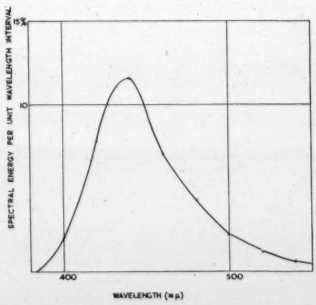


Fig. 8. Fluorescence of an optical bleach

being 365 mμ. The blue fluorescence counteracts the slight yellowness which is nearly always present in white materials, and is sensitively appreciated by the eye as a departure from perfect whiteness. Blue fluorescence can only be excited by the ultra-violet so that such spectacular increases of reflexion,

as occurred with the red pigment, cannot be expected. Filter paper, of reflexion factor 75%, was used as a base and soaked with a solution of the fluorescent material and then dried. A table, similar to the one given in this paper, was then derived from measurements and the increase of reflexion, with sunlight as illuminant, was calculated and is shown in Fig. 9. The conditions of this experiment are doubtless much more favourable than would occur in domestic practice; nevertheless the increase of 25% in reflexion factor is quite considerable.

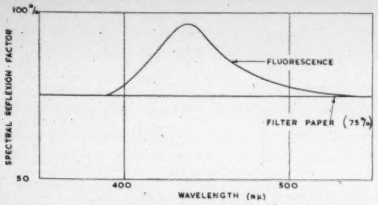


Fig. 9. Spectral reflexion of optical bleach for sunlight

COLORIMETRIC ASPECT

The spectral reflexion table is a complete analysis of a fluorescent material as far as the physical radiation is concerned and the reflected energy for any illuminant can be calculated from it. Both the preparation of the table and the subsequent calculations are lengthy, however, and are rather out of the question for everyday measurement of such colours. If the interest is limited to the colour of the sample only and does not extend to its spectrophotometric curve, then considerable simplification can be made without introducing any approximation. The important property, that the colour can be calculated for any prescribed illuminant, is also retained.

Up to now only spectrophotometric values have been considered. It is possible to replace a complete spectrophotometric curve by three tristimulus values which define the colour of the distribution. The set of spectrophotometric curves, which come from all the excitation wavelengths, can be replaced by three curves of tristimulus values. Each tristimulus value is composed of the sum of two contributions, one from the unmodified reflexion and the other from the fluorescent reflexion. Thus, some twenty curves can be reduced to three. The curve of tristimulus value X is shown in Fig. 10. The incident light of equal energy distribution is represented by the usual $\bar{x}$ distribution curve. The unmodified reflexion is shown as a proportion of this according to the reflexion factor. The tristimulus value X of the fluorescent light, when added to the ordinary reflexion, makes up the total reflexion. Thus, for example, at wavelength 580 mμ, the tristimulus value of the incident light is 1·46. The reflexion factor is 8·0% so the tristimulus value X for the unmodified reflexion is 0·12. The fluorescence, as shown in Fig. 3, has a tristimulus value X equal to 0·41. The tristimulus value of the total reflexion is therefore 0·53. As can be seen from the figure, tristimulus value X is present in the reflected light, in the ultra-violet and in the blue-green, but is absent in the incident light. The total tristimulus value X for an illuminant having equal energy is given by the area under the curve. For any other illuminant the appropriate tristimulus curve is obtained by multiplying by the energy curve in the usual way and the value for that illuminant is the area under the curve. Similar curves may be obtained for the Y and Z tristimulus values.

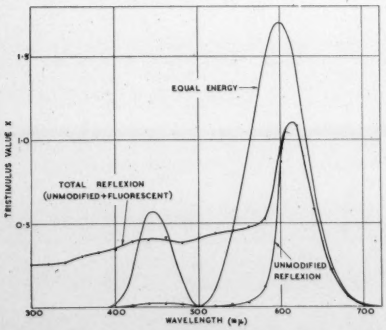


Fig. 10. Curve of tristimulus value X for fluorescent ink (red)

Thus, this group of *three* tristimulus curves allows the colour to be calculated for any prescribed illuminant. With an ordinary non-fluorescent sample we have a measured reflexion curve and three distribution curves. Here we have three measured curves of tristimulus values and one energy curve. The amount of calculation is therefore just the same in both cases. It is to be noted that the use of tristimulus values does not introduce any approximation, and can also deal with fluorescence of any spectral distribution.

The three tristimulus curves may be found either by calculation from the general table or they can be measured directly. The second monochromator, which also functions as a template colorimeter, can be used to obtain the tristimulus value of the reflected monochromatic and the fluorescent light together in one reading. This shortens the time required for the complete measurement considerably. A separate monochromator, however, is still necessary to illuminate the sample. It must be admitted that the measurement of fluorescent specimens is still rather an elaborate proceeding. The difficulties are confined to the practice. On the theoretical side, the machinery of the C.I.E. system shows itself quite capable of including fluorescent colours in its scope, although the system was not expressly formulated to include them.

ACKNOWLEDGEMENT

The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

REFERENCES

(1) Reimann, G., Judd, D. B., and Keegan, H. J. *J. Opt. Soc. Amer.*, **36**, p. 128 (1946).
(2) Tyler, J. E., and Callahan, F. P. *J. Opt. Soc. Amer.*, **41**, p. 997 (1951).
(3) Walsh, A. *J. Opt. Soc. Amer.*, **42**, p. 95 (1952).
(4) Donaldson, R. *J. Sci. Instrum.*, **29**, p. 150 (1952).
(5) Wawilow, S. I. *Phil. Mag.*, **43**, p. 307 (1922).
(6) Bowen, E. J. *Proc. Roy. Soc. A*, **154**, p. 349 (1936).
(7) Johnson, F. S., Watanabe, K., and Tousey, R. *J. Opt. Soc. Amer.*, **41**, p. 702 (1951).
(8) Bowen, E. J., and Sawtell, J. W. *Trans Faraday Soc.*, **33**, p. 1425 (1937).
(9) Moon, P. *J. Franklin Inst.*, **230**, p. 604 (1940).

Direct-reading metal spectroscopy with a d.c. arc

By S. C. Baker, M.Sc., A.Inst.P., Newcastle Technical College, New South Wales, Australia

[*Paper first received* 27 *April, and in final form* 24 *August*, 1953]

A d.c. globule arc that hisses vigorously has been developed in conjunction with a direct-reading spectrometer built on a wavelength spectrometer and measuring intensity ratios by the null indication of a symmetrical cathode-follower bridge circuit. Self-electrodes are held in massive brass supports with the anode uppermost and the arc run until large molten globules form and stabilize on both electrodes. The relationship between the manganese concentration in carbon steels and the intensity ratio of manganese 4754 Å to iron 5455·6 Å is linear in the range 0·3% to 1·3% manganese, and standard deviations of 2% of content have been obtained. Standard deviations of 4% have been obtained by ordinary photographic photometry using manganese 4754 Å and iron 4707·281 Å. A method of preparing a working curve from a knowledge of the background intensity of the manganese line and a single standard sample is suggested. There is a discontinuity in the volt-ampere characteristic of the arc, but the curve is nearly parallel to the current axis at the operating point.

This paper is the result of attempts to estimate manganese in carbon steels with simpler and cheaper apparatus than commercially available quantometers. Several published circuits that measured spectrum line intensity ratios with a pair of photomultiplier tubes attached to a suitable spectrograph[1,2,3] were tried but found unsuitable, and a new bridge circuit has been evolved. The ordinary d.c. arc and a wavelength spectrometer were used as matters of convenience and availability in the first instance, but when suitably modified they have proved satisfactory for the present purpose. Visible lines are used throughout and the attachments are therefore being fitted to a "Spekker" steeloscope (by Hilger and Watts Ltd.) for works operation.

OPTICAL DETAILS

The arc is focused on a wavelength spectrometer (by Hilger and Watts Ltd.) by the intermediate method of slit illumination and the selected pair of spectrum lines presented to the photomultiplier tubes as represented in Fig. 1. Exit slits A and A', 0·07 mm wide by 5 mm high, are supported in the position normally occupied by the photographic plate by accurately-machined brass ways that provide universal adjustments. The entrance slit of the spectrograph is 0·04 mm wide by 2 mm high. Achromatic lenses L and L' of focal length 5·5 cm and diameter 1·5 cm focus images of the exit slits onto ground glass diffusing screens G and G' respectively, supported about 1 cm in front of the glass envelopes of the photomultiplier tubes P and P'. These diffusing screens and lenses minimize the effects of arc wander and render the positioning of the photomultiplier tubes less critical.[3] The slit and lens fit into a light-tight tube of telescopic construction so that, after first locating the selected spectrum line on the cross-wire of an auxiliary telescope, the components can be added and adjusted in turn. The photomultiplier tube with attached diffuser is separately adjustable and is the last unit placed in position. The telescopic tubes support all accessories directly on the plateholder of the spectrograph and, at the same time, all extraneous light is excluded without having to enclose the whole unit in a light-tight box.

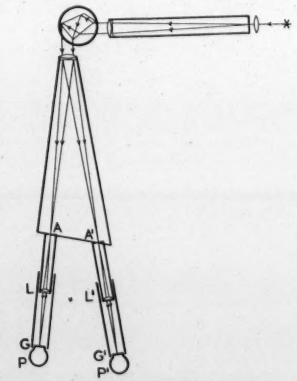


Fig. 1. Plan of wavelength spectrometer

Exit slits AA', convex lenses LL', glass diffusers GG' and photomultiplier tubes PP' attached.

ELECTRONIC EQUIPMENT

A simplified diagram of the amplifier bridge circuit is given in Fig. 2. It is described in detail elsewhere.[4] A pair of photomultiplier tubes, operated from a conventional regulated power supply, generate the photocurrents i_1 and i_2, which independently charge the storage condensers C_1 and C_2. The load visitors R_1 and R_2 limit the potentials built up on the condensers to the mean values R_1, i_1 and R_2, i_2. The triodes, V_1 and V_2, operated from an ordinary unregulated power supply, are interposed as cathode-follower impedance

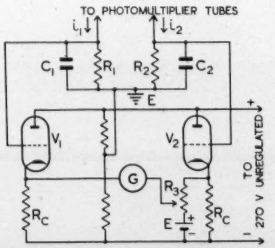


Fig. 2. Simplified diagram of the amplifier bridge circuit

converters between each condenser and each side of the measuring potentiometer-galvanometer circuit. An accurately-wound potentiometer R_3 and short-period galvanometer G are associated with the low-impedance output points of the cathode-followers. This potentiometer taps off any desired fraction of the cathode signal potential to V_2 and is adjusted to yield a null indication on G. The voltage source E backs off the no-signal static component of anode current from R_3. Complete compensation for photomultiplier tube dark currents is provided, and the circuit is virtually independent of supply potential fluctuations and is free from drift. A switching scheme enables adjustments such that the galvanometer remains at zero for all positions of the potentiometer prior to an exposure.

If r denotes the tapped-off portion of R_3, and R its total resistance, it can be shown[4] that

$$i_1/i_2 = r/R \qquad (1)$$

for a balance so that the ratio of the mean values of the photocurrents is proportional to the potentiometer scale reading. For a given ratio, this reading is independent of the absolute values of the currents over a much wider range than those encountered here, the one restriction being that the greater photocurrent must be applied to the grid of V_2 to obtain a null point on the potentiometer scale. A calibration curve for the electronic equipment enables the conversion of nominal scale readings to true intensity ratios.

The control exercised by the potentiometer over the galvanometer is virtually instantaneous and a bridge balance is rapidly and easily obtained. Providing the ratio of the mean levels of the photocurrents remains constant, the null point is independent of the time for which these currents have been flowing. It is advisable, however, to delay the final setting until a period at least equal to the 15 sec time constant of the input circuits has elapsed, for then the sensitivity of the equipment is approaching a maximum.[4]

ANALYTICAL RESULTS

Steel electrodes varying between 5 and 6 mm in diameter are clamped axially in fluted brass cylinders of diameter 3·8 cm and height 2·5 cm, the electrode tips protruding 0·5 cm out of the ends of the cylinders. Power is supplied by a 140 V storage battery in series with a 25 Ω rheostat, the arc gap being adjusted to give a steady current of 4·25 A. With the anode uppermost, large molten globules form on the ends of both electrodes in about one minute and then the intensity ratio of certain spectrum lines settles down to a steady value during the next minute. The arc now emits a vigorous hissing sound and an "exposure" of 15 sec can be made at any time. Careful shaping of the electrodes has no special merit, but cleaning them shortens the stabilizing time if they are oxidized. (It was observed, however, that if the globules were not removed after a previous run, the same steady state could be attained on a re-run within a few seconds.)

After this arc has stabilized, fluctuations in the photomultiplier tube currents are only slightly greater than those due to dark-currents alone. Furthermore, the periods of these fluctuations are half a second or less and, consequently, reasonably accurate potentiometer adjustments can be made with the storage condensers disconnected. Nevertheless, maximum accuracy and ease of operation are obtained by utilizing the averaging effect of the condensers and delaying the opening of the light shutter until the globules are well formed on the electrodes. After opening the shutter, as the storage condensers gradually charge, the galvanometer needle is returned to zero by adjustment of the potentiometer, a final adjustment being delayed until complete stability of the arc is attained. This is readily observed by noting when galvanometer drift in either direction has ceased, but with experience it is usually possible to judge when the stable state has been reached by direct observation of the arc itself. The total time for a single determination is about 2 min.

To determine the nature of the working curve in the range 0·3% to 1·3% manganese content, six analysed pairs of electrodes were used. The mean potentiometer scale readings for six runs of each sample are plotted against the nominal alloy content at A in Fig. 3. The maximum deviation from

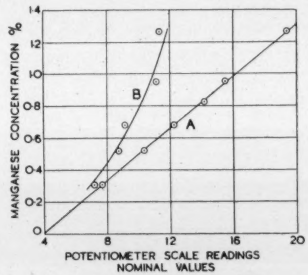


Fig. 3. Working curves for manganese

A, hissing arc; B, silent arc.

the straight line is of the order of 0·007% manganese. Precision data obtained from more than thirty determinations of each sample are given in Table 1. It will be observed that the standard deviation increases with increasing manganese concentration, but there is a tendency for it to decrease when expressed as a percentage of the amount present. These results were obtained with one photomultiplier tube receiving the 4754·0 Å manganese line and the other receiving the 5455·06 Å iron line, which is at a convenient distance from the manganese line.

Table 1. *Electronic measurements*

Manganese concentration %	*Standard deviation % Mn*	*Standard deviation % of Mn content*
0·31	0·008	2·6
0·52	0·012	2·3
0·68	0·014	2·1
0·83	0·016	1·9
0·96	0·014	1·5
1·27	0·019	1·5

A smaller number of observations was made upon the manganese lines 4783·4 Å and 4823·5 Å but the accuracy was lower than that quoted above. Also the hissing sound could be nearly eliminated by increasing the resistance in the arc circuit but satisfactory analytical results could not be obtained with the silent arc. Such readings as could be estimated with the 4754·0 Å manganese line suggest the curve B in Fig. 3.

BACKGROUND EFFECT

By reference to the calibration curve of the electronic bridge circuit the working curve A in Fig. 3 has been re-drawn at A in Fig. 4 to show manganese concentration *versus* the actual ratio of the photomultiplier tube currents. Denoting the luminous fluxes of the manganese and iron lines passing through the exit slits by I_m and I_f respectively and the corresponding background fluxes by B_m and B_f we have from the equation (1)

$$(I_m + B_m/I_f + B_f) = k(r/R) \qquad (2)$$

where k is a proportionality constant and r/R is the ratio of the photomultiplier tube currents. If $(I_f + B_f)$ is maintained constant this reduces to

$$I_m = k'r/R - B_m \qquad (3)$$

Assuming that the light flux from the manganese line is proportional to the concentration Q of manganese in the arc equation (3) can be rewritten

$$Q = Kr - b \qquad (4)$$

where K and b are constants. Evidently then, the intercept b on the concentration axis in Fig. 4 measures the background

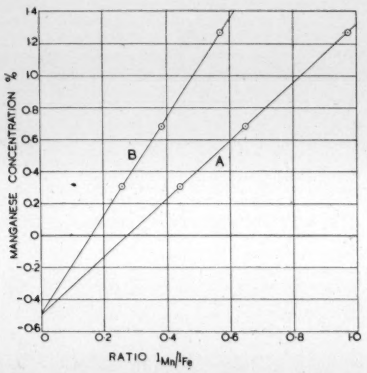


Fig. 4. Working curves for manganese with photocurrent ratios as abscissa showing constant background effect and alteration of calibration with change in equipment constants

flux through the manganese line exit slit as an equivalent percentage of manganese, and this depends only on the spectrometer entrance and manganese line exit slit widths. On the other hand the constant K would depend on many factors, such as the photomultiplier tube luminous sensitivity, alinement of photocathodes to spectrum line images and iron line exit slit width, as well as upon the calibration of the measuring circuit. Any alteration to any of these factors should therefore alter the slope of the working curve but not the concentration intercept. This feature was tested by upsetting the overall calibration of the equipment by the simple expedient of removing the diffuser glass from the photomultiplier tube receiving the iron line. The curve B in Fig. 4 was then plotted from a single determination on each of three standard samples; it intercepts the concentration axis within 0·02% of curve A. This suggests the possibility of preparing a working curve by arcing a single standard sample and joining the point so plotted to the fixed intercept point.

ACCURACY

The results so far obtained are only slightly inferior to those claimed for much more complicated and costly equipment and the accuracy is ample for the immediate application. Apart from the light source, the errors are largely due to the sensitivity of the optical equipment to mechanical vibrations and temperature changes, but it is believed this will be overcome by a sturdier method of attaching A, L, G and P (Fig. 1) to the Steeloscope. Furthermore, the curvature of the spectrum lines restricts the choice of slit widths, but this can probably be eliminated by using a curved entrance slit. The relatively wide exit slits used increase the background flux and although this does not upset the linearity of the working curve it restricts the operating range of the measuring potentiometer by displacing the curve horizontally. For example, Fig. 4 shows that at 0·5% manganese concentration, approximately half the light passing through the manganese exit slit is due to background. This difficulty could be minimized by one of the usual spectroscopic methods[5] but it is desired to avoid this further complication for works operation. Separate tests show that the electronic equipment measures the line intensity ratio with an error less than 0·5%.

PHOTOMETRIC INVESTIGATION

The reproducibility of the hissing arc was examined by preparing a working curve by ordinary photographic photometry, using a Hilger E185 spectrograph with 30° and 60° glass prisms and achromatic Littrow lens. A slit width of 0·17 mm and the intermediate method of slit illumination were used in conjunction with the same steel electrodes. Ilford thin-film half-tone plates were exposed for 15 sec to the stabilized arc. The spectrograms were measured with a Leeds and Northrup recording microphotometer, but the 4707·281 Å iron line was used as internal standard because it is close to the 4750 Å manganese line. The working curve is linear and has a slope somewhat less than that of A in Fig. 4. Standard deviations based on twenty determinations of each sample are given in Table 2. Since the densities of

Table 2. *Photographic measurements*

Manganese concentration %	*Standard deviation % Mn*	*Standard deviation % of Mn content*
0·31	0·016	5·2
0·52	0·017	3·3
0·68	0·019	3·3
0·83	0·036	4·3
0·96	0·041	4·3
1·27	0·045	3·6

the manganese lines from the first and last steels fall at the extreme limits of the plate characteristic curve, these results indicate that the reproducibility of this arc is remarkably good. It must therefore be kept in mind that at present the investigation is restricted to the one range of manganese in carbon steels. A few measurements with the cathode uppermost gave standard deviations 1% to 2% greater than those in Table 2.

PHYSICAL FEATURES OF THE ARC

Before the electrodes were held in the brass cylinders it was observed that globules could be formed on both electrodes quickest by having the anode uppermost, because then gravity

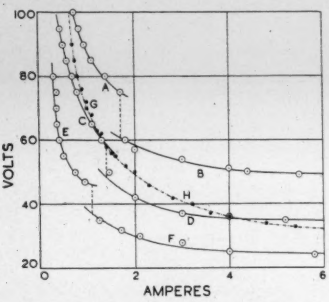


Fig. 5. Volt-ampere curves of d.c. arc

Steel electrodes: *AB*, 6 mm gap; *CD*, 3 mm gap; *EF*, 1 mm gap; *A*, *C* and *E*, silent arc; *B*, *D* and *F*, hissing arc; *GH*, copper electrodes and 3 mm gap.

aided the transfer of a globule from the top to the lower electrode when they were brought together. Subsequently the cylinders were added as cooling vanes to reduce the rate at which the electrodes burnt away during long spectroscopic adjustments. The net effect of the cylinders is that globules form quickly on both electrodes, without bringing them together again, after first striking the arc, and the arc remains steady for longer periods. The flame surrounding the arc plasma is wider at the top than at the bottom and the general appearance of the arc gives the impression that this is associated with convection currents. A large inductance in the arc supply line lengthens the time required for the arc to stabilize.

Characteristic volt-ampere curves of this arc are plotted in Fig. 5 for carbon steel electrodes 6 mm in diameter. The curve *AB* is for a 6 mm arc gap, *CD* for a 3 mm gap and *EF* for a 1 mm gap. The curves are extrapolated slightly on either side of a discontinuity represented by the broken lines, because the arc can be run for short periods in a kind of meta-stable state. The general form of the curves is similar to those published by Ayrton[6] for the carbon arc, and by analogy the parts *A*, *C* and *E* of the volt-ampere curves are here described as the silent arc and the parts *B*, *D* and *F* as the hissing arc. The discontinuity does not occur when copper or tungsten electrodes are arced under the same conditions, and it appears that this arc differs from Milbourn's copper globule arc.[7] The curve *GH* in Fig. 5 is a volt-ampere characteristic for copper electrodes with a 3 mm arc gap and, whilst Nottingham's[8] and Ayrton's[6] formulae fit this curve quite well, neither formula suits the steel curves.

An arc current of 4·25 A was used throughout the analytical work and this necessitated an arc gap of 3 mm between the tips of the glowing globules on the electrodes. Evidently the voltage drop across the arc is very steady because the operating point is at the right of *D* in Fig. 5 where the curve is nearly parallel to the current axis.

An interesting feature of the silent steel arc is the appearance of some of the second positive nitrogen bands in a narrow region near the anode. These bands, together with the cathode layer effect and some reversals, can be seen in Fig. 6, which was prepared by focusing a 3 mm arc and both globule tips on to a spectrograph slit 9 mm long. Evidently in this case the arc current does not evaporate sufficient metal to eliminate all air lines, and the transition from the silent to the hissing state is due to the replacement of air molecules by metallic atoms.

A cathode-ray oscilloscope connected across the hissing arc reveals the presence of audio-frequency oscillations of small amplitude in the arc current. With a high-sweep frequency it can be seen that the phases, frequencies and amplitudes of these oscillations vary irregularly, but the oscillogram at the top of Fig. 7 had to be taken with a time base of 15 msec and camera shutter speed of 1/250 sec to secure a recording with available equipment. The lower oscillogram was produced by a crystal microphone held near the arc and connected to the second channel of the oscilloscope. Identical oscillograms are observed with the silent arc, but the amplitudes are much smaller, a sudden change of amplitude coinciding with the discontinuity in the volt-ampere curve. Evidently, then, the hissing sound is closely

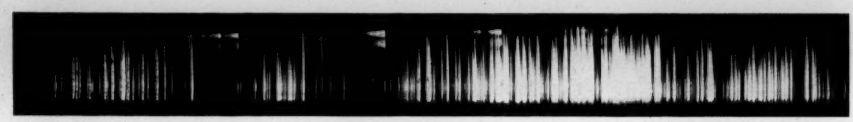


Fig. 6. Spectrogram of silent steel arc in the range 3000 Å to 4000 Å showing second positive nitrogen bands near the anode, reversals and the cathode layer effect

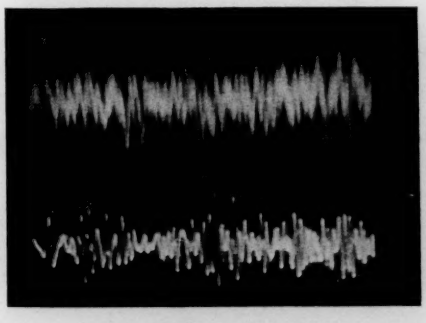


Fig. 7. Double beam oscillogram

Top: oscillations of hissing arc current.
Bottom: sound detected by crystal microphone.

related to minute and irregular fluctuations of the arc current.

It is suggested that the oscillations are due to excessive evaporation from the glowing globules at the high-current density required by a short path and relatively high-potential difference between the electrodes. The resulting cloud of vapour cools the arc, thereby reducing, in particular, the anode fall of potential so that the arc current increases. This takes the operating conditions momentarily to an unstable point on the load-line curve from which it returns to the stable point and the cycle repeats. The oscillograms indicate that several such oscillations occur simultaneously, possibly each in a separate arc streamer, resulting in the unpleasant hissing noise that is quite distinct from the pure tone of a singing arc.

LINE CLASSIFICATION

The Kayser–Ritschl tables designate the manganese lines at 4754·0 Å, 4783·42 Å and 4823·5 Å as MnI and the iron lines at 5455·6 Å and 4707·281 Å as FeI, whereas Harrison's M.I.T. tables leave these manganese lines unclassified and list only the iron lines as normal arc lines. In the present work the manganese 4754 Å line is evidently homologous with the FeI lines, so this manganese line is a normal arc line in the hissing arc.

ACKNOWLEDGEMENTS

This investigation was commenced by Mr. J. Newburn and I am indebted to him and to Mr. B. Gwynne of Rylands Bros. (Aust.) Pty. Ltd. for the analysed steel electrodes. The electronic bridge was brought into satisfactory operation by Mr. W. G. Walker, who also made the electronic measurements. The photographic measurements and all mechanical parts were made by Mr. E. F. Palmer.

REFERENCES

(1) BOETTNER, E. A., and BREWINGTON, G. P. *J. Opt. Soc. Amer.*, **34**, p. 6 (1944).
(2) NAHSTOLL, G. A., and BRYAN, F. R. *J. Opt. Soc. Amer.*, **35**, p. 10 (1945).
(3) SAUNDERSON, J. L., CALDECOURT, V. J., and PETERSON, E. W. *J. Opt. Soc. Amer.*, **35**, p. 11 (1945).
(4) WALKER, W. G., and BAKER, S. C. *J. Sci. Instrum.*, **30**, p. 328 (1953).
(5) TWYMAN, F. *Metal Spectroscopy* (London: Charles Griffin and Co. Ltd., 1951).
(6) AYRTON, H. *The Electric Arc* (London: The "Electrician" Printing and Publishing Co., 1902).
(7) MILBOURN, M. *J. Inst. Met.*, **69**, p. 441 (1943).
(8) NOTTINGHAM, W. B. *Phys. Rev.*, **37**, p. 471 (1926).

The specific heat of polymerized Araldite and Wood's metal between 1·5 and 20°K

By D. H. PARKINSON, M.A., D.Phil., and J. E. QUARRINGTON, Radar Research Establishment, Great Malvern, Worcs.

[*Paper, received* 22 *February*, 1954]

Araldite varnish and Wood's metal are frequently used in the construction of low temperature apparatus. Their heat capacities have been determined in the temperature range 1·5 to 20° K and over most of it, weight for weight, they are found to be very much higher than those of copper.

In the construction of apparatus used in low temperature experiments, varnishes and synthetic resins are frequently used for the electrical insulation of resistance thermometer windings, and also for providing thermal contact between such coils and metallic parts of the apparatus. Of the various materials available for these purposes, Araldite Type 1 has been found highly satisfactory. The makers (Aero Products, Ltd.) specify baking for 1 hour at 180 to 190° C. Treated in this way Araldite shows excellent bonding properties to metals and also to glass and other substances. In many experiments where this and similar varnishes are used, it is highly desirable to know the heat capacity involved, for it may well become several times that of the metallic parts of the apparatus at temperatures in the liquid helium range. To this end the heat capacity of Araldite Type 1 has been measured between 1·5 and 20° K.

Of the commercially available solders, Wood's metal is very useful for making the final vacuum seal in apparatus such as calorimeters. This material, like Araldite, has a somewhat high heat capacity at low temperatures and it also has been measured between 1·5 and 20° K.

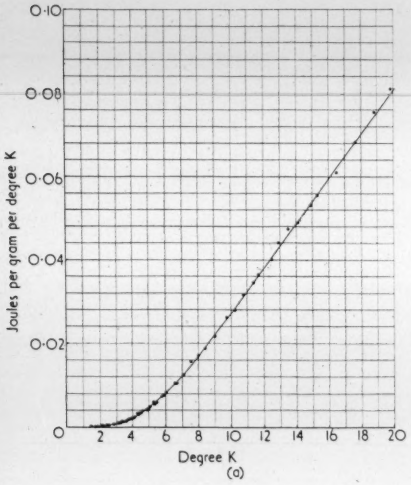
EXPERIMENTAL

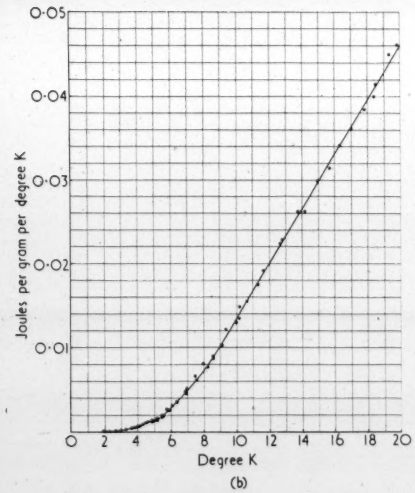
The measurements were carried out in a two-stage calorimeter similar to that described by Hill.(1) Thermometer and heater were wound on the outside and held in place by Araldite varnish. The thermometers were of leaded brass wire between 1 and 4° K, phosphor bronze from 3·0 to 5·5° K, and constantan from 5 to 20° K. A Linde liquefier-cryostat combination, to be described elsewhere,(2) was used. The method of experiment has already been adequately discussed by various authors,(1,2,3) cooling being brought about by condensation of helium into a calorimeter chamber.

The Araldite specimen was prepared according to the makers' specifications in layers about 1·5 mm thick on metal sheets, from which it was chipped in flakes of about 1 to 2 cm^2 area. The total weight used was 33·34 g.

The Wood's metal was supplied by Judex Chemicals and had the following composition: tin 12·5%, cadmium 12·5%, lead 25%, bismuth 50%. Its melting point was given as 68° C. Sticks of this, approximately 1 cm in diameter and 4 cm long were scraped clean before the measurements, the total weight used being 175 g.

In both cases the specimens were sealed into the calorimeter with approximately 2 mm pressure of helium thermal exchange gas. The total heat capacity of the calorimeter had already been determined and was nowhere greater than 10% of that of the specimen.





Joules per gram per degree K *versus* degrees K

(*a*) Araldite.
(*b*) Wood's metal.

RESULTS

The experimental results are plotted in the figure; there are sufficient observations for the best curve to be drawn by eye quite accurately. Experimental errors are estimated at approximately 1% throughout. With Wood's metal there is evidence of a superconducting transition at 4·8° K.

The smoothed specific heat values are given in the table together with those for copper.[4]

Specific heats of Araldite and Wood's metal

Units 10^{-4} *j/g*

T° K	*Araldite Type 1*	*Wood's metal*	*Copper*
1·5	0·6	0·2	0·2
2·0	2·4	0·6	0·3
2·5	5·0	1·2	0·42
3·0	8·9	2·4	0·57
3·5	14·6	3·9	0·75
4·0	22·5	6·2	0·98
5·0	46	13·9	1·6
6·0	82	29	2·5
7·0	124	50·5	3·6
8·0	169	76	5·1
9·0	218	104	7·0
10·0	272	134	9·3
11·0	325	165	12·1
12·0	379	198	15·4
13·0	433	231	19·3
14·0	488	264	23·9
15·0	542	297	29·2
16·0	597	329	35·2
17·0	651	361	41·7
18·0	704	394	49·4
19·0	758	427	57·8
20·0	811	460	67·3

It is clear that except at the lowest temperatures these heat capacities are very high compared with those of copper. Where calorimetry is envisaged with materials of low specific heat it is therefore quite possible for the heat capacity of only a fraction of a gram of these substances, used for, say, holding a thermometer in place, to be equal to that of many grams of the substance under investigation.

ACKNOWLEDGEMENTS

The authors wish to thank Miss M. Davies for her assistance in these experiments and with the calculations involved, and Mr. W. A. Hoyes and others of the staff who produced the liquid hydrogen. They are also indebted to the Chief Scientist, Ministry of Supply, and to the Controller, H.M. Stationery Office, for permission to publish this paper. Crown copyright is reserved.

REFERENCES

(1) HILL, R. W. *J. Sci. Instrum.*, **30**, p. 331 (1953).

(2) PARKINSON, D. H. *J. Sci. Instrum.*, **31**, p. 178 (1954).

(3) PICKARD, G. L., and SIMON, F. E. *Proc. Phys. Soc. [London]*, **61**, p. 1 (1948).

(4) KOK, J. A., and KEESOM, W. H. *Physica*, **3**, p. 1035 (1936).

Densitometric evaluation of coal-dust stains on filter paper

By J. G. Dawes, Ph.D., A.M.I.Min.E., F.Inst.P., Safety in Mines Research Establishment, Sheffield

[*Paper first received* 16 *June*, 1952, *and in final form* 1 *September*, 1953]

The theory of Davies and Aylward[1] for interpreting the light-screening power of deposits of coal-dust on a uniform filter paper is extended, and the modified theory is tested using data from samples taken in a dust tunnel.

Davies and Aylward[1] discuss the photoelectric measurement of coal-dust stains on filter paper, and present formulae for interpreting the light-screening power of deposits of coal-dust collected on different types of filter paper. The formula used for interpreting the stains collected on a uniform paper, e.g. paper No. 3 (Davies and Aylward[1]) is of particular interest; this formula may be expressed as

$$i/i_0 = \exp(-\alpha\nu_T) \qquad (1)$$

where ν_T is the total number of coal-dust particles collected; α is a parameter dependent on the size distribution of the particles collected and $\alpha\nu_T$ is the total projected area of the particles collected; i_0 is the current through the photoelectric cell when the paper is clean, and i is the current when the paper is stained by the coal-dust deposit.

The authors suggest, from comparatively few results, that the above formula is adequate for the two finest clouds used in their tests, but that an empirical correction (namely the argument of the exponential function must be multiplied by 1·2) must be made in order to match the formula to the results for the two coarser clouds.

The present paper gives new data on which the theory of Davies and Aylward can be tested, and proposes a development of their theory to account for the discrepancies observed.

EXPERIMENTAL ARRANGEMENTS

The tests were carried out in the Sheffield (S.M.R.E.) dust tunnel. The cross-section of the tunnel is 1 ft square, and coal-dust is continuously fed into the air-stream driven through the tunnel by a fan. The tests were carried out with an air velocity of 100 ft/min in the tunnel (with one exception at 200 ft/min). The coal-dust is projected into the air-stream by a special dust disperser and the aerosol is discharged into the atmosphere on reaching the open end of the tunnel. The sampling position is some 28 ft from the dust source. The dust cloud at the sampling position was always monitored during a given run by a thermal precipitator[2] and usually by a Pneumoconiosis Research Unit (P.R.U.) handpump,[3] which at frequent intervals took a sample, always with the same number of pump strokes. The consistency of the concentration on a given run and of the size distribution in the size range less than 10 μ on a given series of runs was established by means of the thermal precipitator, and overall by the P.R.U. handpump taking a constant number of strokes for each sample.

The samples taken for the purpose of test were taken with a handpump, the number of pump strokes being varied over a wide range so that a range of stain densities could be obtained. In the diagrams given below each plotted point is the mean of four observations with the test handpump taking a given number of strokes.

Although the size distribution in the size range 0·5 to 10·0 μ could be determined by thermal precipitator sampling with relative precision, the size distribution above 10 μ could not be determined accurately. That it remained relatively constant during a given series was deduced from the monitor handpump observations and from internal consistency in the data presented below.

Since the size distribution in a given experiment was constant, the usual parameter,[3] namely the count of particles per cm^3 in the size range 0·5 to 5·0 $\mu = \nu_R$, was taken as descriptive of the cloud concentration. The total area collected, i.e. $\alpha\nu_T = \beta\nu_R$ where β is constant for a given size distribution. Handpump samples involving two to one hundred pump strokes were collected and for each sample the corresponding values of i_0 and i were recorded, these being measured on a P.R.U. densitometer.[3] The size distribution of the dust collected on the filter paper was varied in different series of experiments as follows. Most of the experiments were carried out with a P.R.U. handpump sampling in a cloud obtained by dispersing Silkstone coal-dust, ground so that 85% passed a 240 B.S. mesh. This dust is prepared, by a standard routine method, for coal-dust explosion trials at the Buxton station of the Safety in Mines Research Establishment. The size distribution of this dust when airborne, as estimated microscopically from thermal precipitator samples, is, in the size range 0·5 to 10 μ, described to a first approximation by $dF_N/dD = 0{\cdot}8\exp(-0{\cdot}8D)$, where dF_N/dD is the number frequency function and D the microscope particle size in microns (see Wynn and Dawes[4]).

The second method was to use the same Silkstone coal-dust, but to collect the sample on the filter paper, not with a P.R.U. handpump, but with a handpump fitted with a size selector designed to give a 50% collection efficiency for 5 μ spheres of unit density.

The third method of experiment was to use a P.R.U. handpump sampling, not in the cloud of Silkstone coal-dust, but in a cloud obtained by dispersing the fraction of dust, collected underground during a survey on the underbelt deposition of coal-dust from a conveyor system, which passed a 240 B.S. mesh. This dust, when dispersed in the tunnel and sampled by a thermal precipitator, has, in the size range 0·5–10 μ microscope diameter, a size distribution not dissimilar to the Silkstone coal-dust, but there are proportionally more coarse particles present. Examination of the dust reveals that there is a wider range of shape factor distribution in this dust than there is in the mill-ground Silkstone, especially in the coarser fraction. Moreover, the trials with this dust were carried out at an air velocity of 200 ft/min in the wind tunnel, so that there was less deposition of the coarse dust between the dust cloud former and the sampling position.

TEST OF THE THEORY OF DAVIES AND AYLWARD

Using the above techniques, the theory of Davies and Aylward[1] was tested using two different filter papers. The first paper was of the same type as that used by Davies and Aylward and called by them paper No. 3. The second paper is of a similar type, being of uniform texture and made from esparto fibre by another manufacturer. This paper will be called paper WT.2.

The relations between i/i_0 and ν_R for these two papers are shown in Figs. 1 and 2, where log/linear grids are used.

If equation (1) is valid, the plots should be straight lines. In practice, they are significantly curvilinear. Moreover, simple multiplication of the argument of the exponential by a constant factor will not alter the curvilinearity. It is apparent that the theory must be refined, and the following theory has been developed.

THEORY OF LIGHT SCREENING BY OPAQUE PARTICLES ON A UNIFORM PAPER

Let the initial photocell current for a clean paper $= i_0$ and the final photocell current using a stained paper $= i$. It is assumed, with Davies and Aylward, that few particles less than $0{\cdot}5\ \mu$ in size are collected and that we are dealing with simple obscuration effects.

Then the area obscured $\propto (i_0 - i) = \lambda(i_0 - i)$ the corresponding initial area being λi_0, where λ is a constant. Hence, per unit area of the paper, the projected area of the dust collected $= (1 - i/i_0) = f(\nu)$ where ν is the number of particles collected per unit area of the paper.

Suppose that the number of particles collected per unit area of the paper is increased from ν to $\nu + \delta\nu$. The added projected area is then $\alpha\delta\nu$ where α is a function of the size distribution of the dust collected on the paper and in a given test is considered to be constant. The proportion of the area of the paper not already obscured $= 1 - f(\nu)$. If there is *equal probability* of the $\delta\nu$ particles falling at any point in a given plane of the paper, then the proportion of the area $\alpha\delta\nu$ falling on clean paper $= 1 - f(\nu)$, i.e. the extra area of dust collected on the clean paper

$$= \alpha\delta\nu[1 - f(\nu)]$$

$$= \text{the decrease of the clean area}$$

$$= -\delta[1 - f(\nu)]$$

i.e. $$\delta[1 - f(\nu)]/[1 - f(\nu)] = -\alpha\delta\nu \qquad (2)$$

Thus $\ln[1 - f(\nu)] = -\alpha\nu + \text{constant}$.
But when $\nu = 0$, $f(\nu) = 0$, i.e. constant $= 0$

i.e. $$1 - f(\nu) = \exp(-\alpha\nu) = i/i_0$$

i.e. $$i/i_0 = \exp(-\alpha\nu) \qquad (3)$$

This is the same result as obtained by Davies and Aylward, who used a different derivation, but also based their result on the assumption of equal probability of a particle falling either on a clean or an obscured part of a plane of the paper. This result, as we have shown (Figs. 1 and 2), is inadequate.

Equation (2) can be rewritten

$$\delta[1 - f(\nu)] = -\alpha\delta\nu[1 - f(\nu)] \qquad (4)$$

and as stated is based on the assumption that there is equal probability that dust will be deposited on obscured and unobscured portions of the cross-section. Dust will not, however, be deposited equally over the plane of the paper, which consists of a matrix of fibres. On the microscopic scale there are preferred zones in the cross-section by which the aerosol will flow and in which the dust will be deposited (Fig. 3). In particular there is an increased probability that dust will be deposited on particles already on a fibre and projecting from it. Instead of equal probability of deposition on obscured and clear parts of a plane in the paper there is an increased probability that dust will be deposited in a position that is already obscured and the probability will be a function of the amount of dust already deposited. The probability that dust will be deposited in a position not already obscured will decay from the limiting value when there is equal probability of dust being deposited in an obscured and an unobscured position [the situation described by equation (4)] and the decay will be a function of the dust

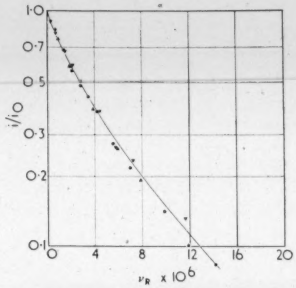


Fig. 1. Test of Davies–Aylward law using paper No. 3. Silkstone coal-dust. P.R.U. handpump

*Concentration**	*No. of strokes*	*Symbol*
4400	2 to 40	●
6200	2 to 32	▲
2300	2 to 64	▼

* Number of particles per cm³ in test cloud, in size range 0.5 to 5.0 μ.

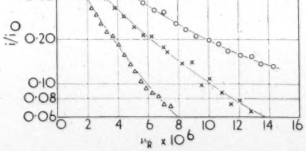


Fig. 2. Test of Davies–Aylward law using paper WT.2

Dust	*Handpump*	*Concentration**	*No. of strokes*	*Symbol*
Ground Silk-stone	S.M.R.E. with size selector.	2680	3 to 67	○
	P.R.U.	2390	3 to 67	×
Mine	P.R.U.	925	5 to 100	△

* Number of particles per cm³ in test cloud, in size range 0.5 to 5.0 μ.

already collected. To account for this effect, an exponential term can be introduced into equation (4) as follows:

$$\delta[1 - f(\nu)] = -\alpha\delta\nu[1 - f(\nu)]\exp(-k_1\alpha\nu) \qquad (5)$$

where k_1 is a parameter characteristic of the paper in use. We then have

$$\ln[1 - f(\nu)] = \frac{1}{k_1}\exp(-k_1\alpha\nu) + \text{constant}$$

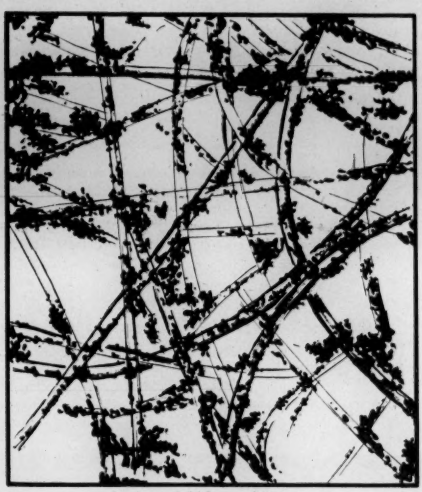


Fig. 3. Coal-dust collected on paper No. 3

when $\nu = 0$, $f(\nu) = 0$

i.e. $$\ln[1 - f(\nu)] = \frac{1}{k_1}\exp(-k_1\alpha\nu) - \frac{1}{k_1} = \ln i/i_0$$

i.e. $$\exp(-k_1\alpha\nu) = 1 + k_1 \ln i/i_0$$

$$-k_1\alpha\nu \log_{10} e = \log_{10}\left(\frac{k_1}{\log_{10} e} \cdot \log_{10} i/i_0 + 1\right)$$

Let $$\log_{10} k_2 = (\log_{10} e)/k_1$$

i.e. $$-\alpha\nu\frac{(\log_{10} e)^2}{\log_{10} k_2} = \log_{10}\left[\frac{\log_{10}(k_2 i/i_0)}{\log_{10} k_2}\right]$$

Writing $$(\log_{10} e)^2/(\log_{10} k_2) = K_1$$

$$\alpha K_1 \nu = \log_{10}\log_{10} k_2 - \log_{10}\log_{10} k_2\, i/i_0$$

or $$\nu_R = K_2[\log_{10}\log_{10} k_2 - \log_{10}\log_{10} k_2\, i/i_0] \qquad (6)$$

where $K_2 = f(\alpha, k_1)$ and is a calibration factor which is constant for a given cloud and a given paper.

We require a value for the paper parameter $k_2 = f(k_1)$. An estimate of k_2 can be obtained using the data on Figs. 1 and 2. The method is to compare observations at two different points on a given estimated curve; thus

$$\frac{\log_{10}\log_{10} k_2 - \log_{10}\log_{10} k_2(i/i_0)_1}{\log_{10}\log_{10} k_2 - \log_{10}\log_{10} k_2(i/i_0)_2} = \frac{(\nu_R)_1}{(\nu_R)_2}$$

where the only unknown is k_2. The values of k_2 obtained are somewhat sensitive to the choice of the best curve through the observations, but for reliable parts of the curve the values range about $k_2 = 100$. For instance, values of k_2 obtained from the data of Fig. 1 are 79, 141, 98, 100, 96, 102 and 100.

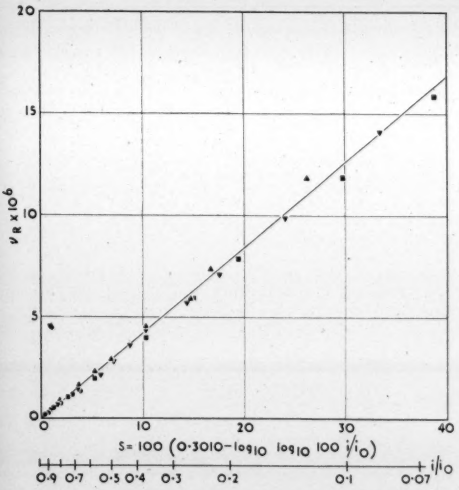


Fig. 4. Test of S scale using paper No. 3. Silkstone coal-dust. P.R.U. handpump.

Cloud	*Symbol*
A	●
B	▼
C	▲
D	■

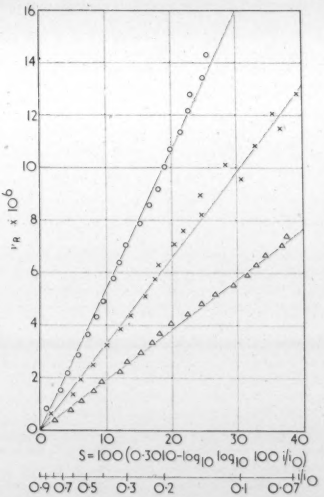


Fig. 5. Test of S scale using paper W.T.2

Dust	*Handpump*	*Symbol*
Ground Silk-stone	S.M.R.E. with size selector	○
	P.R.U.	×
Mine	P.R.U.	△

The convenient value of $k_2 = 100$ has been chosen, and as shown below it is completely adequate within experimental error for a series of different papers of the same general type, and also when the particle size distribution collected on the paper is varied.

For convenience we write

$$S = 100[0{\cdot}3010 - \log_{10}\log_{10} 100\, i/i_0] \quad (7)$$

and

$$v_R = KS \quad (8)$$

where, as before, K is a calibration factor, and is a function of the size distribution of the dust collected on the paper.

TEST OF THE NEW THEORY

Data obtained with paper No. 3 when plotted on the $v_R : S$ co-ordinate system are shown in Fig. 4. Comparable data were obtained in the Silkstone coal-dust cloud with the P.R.U. handpump using two other esparto fibre papers, one made by the manufacturer of paper No. 3 and one an experimental paper made by the firm manufacturing paper W.T.2. Fig. 5 gives data obtained with paper W.T.2 when coal-dust deposits of different size distribution are collected. In no case is there a significant departure from linearity, at least at the 5% probability level. It is, therefore, concluded that the S scale, for the uniform esparto type filter paper, adequately describes the light screening power of collected coal-dust particles. This type of paper is recommended by both Davies and Aylward[1] and by Bourne and Streett[5] for the collection of fine dust from an aerosol. Davies[6] has applied the S scale to the data presented by Davies and Aylward for paper No. 3 and has demonstrated that a significantly better fit at the 5% probability level is obtained than when the simple exponential formula of equation (1) is used, in each of the four dust clouds of different size distribution tested by them.

ACKNOWLEDGEMENTS

The author wishes to express his thanks to members of the staff of the Safety in Mines Research Establishment for assistance in collecting experimental data, to Mr. B. A. Maguire for the statistical tests on linearity and to the Director of the Safety in Mines Research Establishment and the Ministry of Fuel and Power for permission to publish the paper.

REFERENCES

(1) Davies, C. N., and Aylward, M. *Brit. J. Appl. Phys.*, **2**, p. 352 (1951).
(2) Green, H., and Watson, H. H. *M.R.C. Special Report Series No.* 199 (London: H.M.S.O., 1935).
(3) Watson, H. H., and Hounam, R. F. *Colliery Guard*, **176**, p. 447 (1948).
(4) Wynn, A. H. A., and Dawes, J. G. *Safety in Mines Research Establishment Research Report No.* 28 (1951).
(5) Bourne, H. G., and Streett, L. P. *Paper Tr. J.*, **130**, p. 21 (1950).
(6) Davies, C. N. Private communication (1952).

Transverse oscillations in travelling strings

By R. A. Sack, Ph.D., Department of Theoretical Physics, University of Liverpool

[*Paper first received* 23 *November, and in final form* 10 *December*, 1953]

The behaviour of transverse standing waves is examined for a uniform string which is pulled over two smooth supports under constant speed and tension. A harmonic spectrum of resonance frequencies is found which, for small ratios of the speed of travel to the propagation velocity of transverse waves, are only slightly lower than those of a string with clamped ends. The frequency response to a simple harmonic disturbance and the envelope of the oscillations are similar to those of a clamped string, but the phase of oscillation varies from point to point so that the instantaneous configuration of the string is not a sine curve. The conclusions remain valid in the presence of damping.

1. INTRODUCTION

In technical processes in which flexible strings are pulled at high speed over fixed supports, the occurrence of large transverse oscillations is an undesirable but frequent feature. The existence and persistence of such oscillations raises the question to what extent the usual concepts of standing waves, resonance frequencies, nodes, etc., can still be applied to the case of travelling strings.

The problem of small lateral oscillations of an undamped string which travels under constant tension and at constant speed through two fixed points has first been treated by Skutsch.[1] By considering a superposition of two waves running in opposite directions Skutsch calculates the fundamental resonance frequency of the system and derives a formula for determining the configuration of the string at any time when the initial conditions are given. Recently the problem has been treated by Booth and Booth[2] in a similar way.

The method used by Skutsch is not suitable to discuss the behaviour of a string under a sinusoidally alternating influence or to include the effect of damping. In the present paper the steady state response of a travelling string will be determined under the following assumptions, similar to those made in the elementary theory of strings with clamped ends [see Ref. (3)]:

(*a*) the speed of travel v, the mass m per unit length and the tension T of the string are constant throughout;
(*b*) the oscillations are sufficiently small so that only terms of the first order in the deviation from equilibrium need be considered;
(*c*) effects due to gravity, the resistance to bending, and the internal viscosity of the string can be neglected;
(*d*) damping is treated as an effect proportional to the transverse velocity;
(*e*) the string passes over two smooth pegs, a distance L apart, of which one is fixed and the other performs sinusoidal lateral vibrations.

The terms down-stream and up-stream, used below, are self-explanatory.

2. MATHEMATICAL TREATMENT

Forced oscillations without damping. If the position of a point on the string is denoted by x and its transverse deflexion from equilibrium by y, the differential equation describing transverse oscillations is, as in the case of a clamped string,[3] given by

$$\frac{D^2y}{Dt^2} = \frac{T}{m}\frac{\partial^2 y}{\partial x^2} = c^2\frac{\partial^2 y}{\partial x^2} \tag{1}$$

where $c = (T/m)^{\frac{1}{2}}$ is the velocity of propagation of transverse waves; the operator (D/Dt) implies that the co-ordinate system travels with the string. Transformation to the operator $(\partial/\partial t)$ referring to a fixed abscissa means

$$(D/Dt) = (\partial/\partial t) + v(\partial/\partial x) \tag{2}$$

so that equation (1) becomes

$$\frac{\partial^2 y}{\partial t^2} + 2v\frac{\partial^2 y}{\partial t\partial x} - (c^2 - v^2)\frac{\partial^2 y}{\partial x^2} = 0 \tag{3}$$

Let $x = 0$ denote the fixed peg, whereas the peg $x = L$ performs simple harmonic oscillations of frequency $\omega/2\pi$ and of amplitude y_0. Thus

$$y = 0, \quad x = 0 \tag{4}$$

$$y = y_0 \cos \omega t = \tfrac{1}{2}y_0[\exp(i\omega t) + \exp(-i\omega t)], \; x = L \tag{5}$$

We try to express y as a sum of terms of the form $\exp[\pm i(\omega t + \alpha x)]$ or their equivalent. This leads to a quadratic equation in α

$$\omega^2 + 2\omega v\alpha - (c^2 - v^2)\alpha^2 = 0 \tag{6}$$

the solution of which is

$$\alpha = \frac{v\omega}{c^2 - v^2} \pm \frac{c\omega}{c^2 - v^2} = \frac{\omega}{v'} \pm \frac{\omega}{c'} \tag{7}$$

This expression involves two new velocities

$$v' = (c^2 - v^2)/v = T/mv - v \tag{8}$$

and

$$c' = (c^2 - v^2)/c = (T - mv^2)/(Tm)^{\frac{1}{2}} \tag{9}$$

the meanings of which will be explained presently. We can express y as

$$y = \exp[i\omega(t + x/v')][A\sin(\omega x/c') + B\cos(\omega x/c')] + \exp[-i\omega(t + x/v')][A'\sin(\omega x/c') + B'\cos(\omega x/c')] \tag{10}$$

Condition (4) implies that

$$B = B' = 0 \tag{11}$$

and condition (5) yields

$$A = \frac{y_0\exp(-i\omega L/v')}{2\sin(\omega L/c')}, \quad A' = \frac{y_0\exp(i\omega L/v')}{2\sin(\omega L/c')} \tag{12}$$

The final result is thus

$$y = y_0\frac{\sin(\omega x/c')}{\sin(\omega L/c')}\cos\left[\omega\left(t - \frac{L - x}{v'}\right)\right] \tag{13}$$

This result clearly indicates the essential features of the wave set up as a consequence of the vibration of one of the supports. The first factor describes the variation of amplitude along the abscissa. Resonance occurs whenever

$$\omega L/c' = n\pi \; (n \text{ integer}) \tag{14}$$

and the standing wave has a node where

$$\omega x/c' = n'\pi \; (n' \text{ integer}) \tag{15}$$

This factor is entirely equivalent to the corresponding term for the fixed string, and hence the shape of the envelope and the position of the nodes are the same, provided the wave velocity is taken to be c' as defined in equation (9), which can therefore be described as the "mean effective wave velocity." The fundamental frequency of resonance is given by

$$\nu_1 = c'/2L = (c^2 - v^2)/2cL \tag{16}$$

the higher resonance frequencies are integral multiples thereof.

A new feature arises in the second factor of equation (13), which describes the phase of the oscillation at any point. It is seen that the phase is not constant along the string, but that there is a phase shift between any two points proportional to their difference apart such that the vibration at the point upstream lags behind that at the point downstream. The velocity v', defined in equation (8), can be described as a "phase propagation velocity"; it determines the speed at which a given *phase* of oscillation travels upstream; it is infinite for a fixed string.

Forced oscillations with linear damping. Let us assume that the surrounding medium exerts a resistive force on each part of the string which is proportional to the rate of change of shape of the string, though not necessarily to the actual velocity of the material. This means that a section of length dx is subject to a transverse force equal to

$$-\beta m dx(\partial y/\partial t)$$

where β is a constant. Then in equations (1) and (3) a term $-\beta\partial y/\partial t$ must be added on the right-hand side, which leads to

$$\frac{\partial^2 y}{\partial t^2} + \beta\frac{\partial y}{\partial t} + 2v\frac{\partial^2 y}{\partial t\partial x} - (c^2 - v^2)\frac{\partial^2 y}{\partial x^2} = 0 \tag{17}$$

If the boundary conditions (4) and (5) are unaltered, the equation for α corresponding to equation (6) becomes

$$\omega^2 - i\beta\omega + 2\omega v\alpha - (c^2 - v^2)\alpha^2 = 0 \tag{18}$$

with the solution

$$\alpha = \frac{\omega v}{c^2 - v^2} \pm \frac{[c^2\omega^2 - \beta i\omega(c^2 - v^2)]^{\frac{1}{2}}}{c^2 - v^2} = \frac{\omega}{v'} \pm \frac{(\omega^2 - i\beta'\omega)^{\frac{1}{2}}}{c'} \tag{19}$$

where

$$\beta'/\beta = (c^2 - v^2)/c^2 = c'/c \tag{20}$$

It is easy to see without any further calculations that the first term in equation (19) leads to a phase propagation velocity as discussed above; the second term shows that the frequency response, local amplitude of oscillation, position of nodes, etc., are the same as for a fixed string provided c and β are replaced by c' and β' [see also Ref. (3)].

Alternative separation of the wave equation. The original differential equation (1) for undamped oscillations is separable in two variables, one being the time and the other the position measured with respect to the travelling string. For oscillations due to influences which do not move along with the string, this separation is not convenient, and an alternative separation is desirable for which one of the variables is x,

The form of the solution (13) suggests the variables

$$x' = x;\ t' = t + x/v' = t + vx/(c^2 - v^2) \qquad (21)$$

We have

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'};\ \frac{\partial}{\partial x} = \frac{\partial}{\partial x'} + \frac{1}{v'}\frac{\partial}{\partial t'} \qquad (22)$$

so that equation (17) or its factorized form

$$\left\{\left[\frac{\partial}{\partial t} - (c - v)\frac{\partial}{\partial x}\right]\left[\frac{\partial}{\partial t} + (c + v)\frac{\partial}{\partial x}\right] + \beta\frac{\partial}{\partial t}\right\}y = 0 \qquad (23)$$

transforms into

$$\left\{\left[\frac{c}{c + v}\frac{\partial}{\partial t'} - (c - v)\frac{\partial}{\partial x'}\right]\left[\frac{c}{c - v}\frac{\partial}{\partial t'} + (c + v)\frac{\partial}{\partial x'}\right] + \beta\frac{\partial}{\partial t'}\right\}y = 0 \qquad (24)$$

or

$$\frac{c^2}{c^2 - v^2}\frac{\partial^2 y}{\partial t'^2} + (c^2 - v^2)\frac{\partial^2 y}{\partial x'^2} + \beta'\frac{\partial y}{\partial t'} = 0 \qquad (25)$$

which is equivalent to the usual equation for fixed strings with the substitution of c' and β' for c and β according to equation (20). This shows that for free oscillations also the general type of motion is the same for fixed and travelling strings, except for the introduction of the "phase time variable" t' which also depends on position. For example, special methods have to be employed to determine the state of the string at $t > 0$ when both y and $\partial y/\partial t$ are given for all x at $t = 0$. These methods, while not difficult in principle, appear to be rather tedious, and it seems doubtful if this type of problem will be important in practice. For undamped strings, a more convenient method of solving this problem has been given by Skutsch.(1)

3. DISCUSSION

The preceding investigations show the possibility of standing transverse waves along travelling strings; these waves are described by an effective wave velocity c', defined in equation (9), and, in the case of linear damping, by a damping parameter β', defined in equation (20). The formula (16) for the fundamental frequency for the undamped string has already been derived by Skutsch(1) by the simple argument that the effective velocities with respect to a fixed co-ordinate system will be $c + v$ for waves travelling downstream, and $c - v$ for waves travelling upstream; hence any wave after travelling in both directions and being reflected at both supports will revert to its original form after a time τ_1,

$$\tau_1 = \frac{L}{c + v} + \frac{L}{c - v} = \frac{2cL}{c^2 - v^2} = \frac{2L}{c'} = \frac{1}{\nu_1} \qquad (21)$$

and any free oscillation of the (undamped) system is periodic with period τ_1. The mean effective wave velocity c' is the harmonic mean of the effective wave velocities $c + v$ and $c - v$ for waves moving downstream and upstream respectively.

The formulae (9) and (16) show that the longitudinal motion of the string causes a reduction in the resonance frequencies, the relative reduction amounting to v^2/c^2. For small values of v/c, which are primarily of technical importance, this effect is small; if, for example, the frequency response width is 5%, the resonance curve for the travelling string will overlap that of the stationary string even for $v/c = 1/5$. For any investigation which does not take account of the phase of oscillation, such as visual inspection or a photographic exposure extending over several periods, the behaviour of travelling and clamped strings is similar as the shape of the envelope, the position of the nodes, and the variation of amplitude are the same, except for the slight change in the parameters c and β.

The distinguishing feature of the travelling string is the existence of a phase propagation velocity v', defined in equation (8). For forced oscillations the instantaneous shape of the string (as revealed in a flash photograph) is not a sine curve; this is obtained only if the point of observation, e.g. a spot of light, moves upstream with velocity v'. For small values of v/c, v' considerably exceeds c. It should be possible to verify experimentally the existence and value of v' by comparing the phases of oscillation at two or more points along the string.

The effect of gravity, in particular the deviation from a straight line of the shape of a non-oscillating string, moving with constant velocity and under uniform tension, is determined by a parameter γ

$$\gamma = mgL/(T - mv^2) = gL/(c^2 - v^2) \qquad (26)$$

[See also Ref. (4).] As long as γ remains small compared with unity, the deviation is small, and the influence of gravity on oscillations in horizontal planes vanishes and on those in vertical planes can be described by an additional term $-g$ on the right-hand sides of equations (1), (3) and (17).(2) This merely implies that the oscillatory deflexions y calculated in Section 2 have to be measured not from the straight line joining the supports, but from the parabolic shape of the string in dynamic equilibrium. When, however, γ becomes comparable with, or greater than, unity, the linearized equations used above no longer adequately describe the behaviour of the chain, and the solution (13) no longer applies; this will always happen as v approaches c, though this case is hardly of technical importance [see Ref. (2)].

For practical purposes it must be pointed out that for a travelling string the tension and the boundary conditions at the supports will be subject to fluctuations, mainly because of friction, and hence a more erratic behaviour is to be expected than predicted by the above theory. Furthermore, although the values of c' and v' are entirely unaffected by the linear damping coefficient β, more complicated laws of damping may modify them somewhat. The general conclusions, however, relating to standing waves and the phase propagation velocity should apply even in such cases.

ACKNOWLEDGEMENTS

This work formed part of the technological research programme of the British Rayon Research Association, Manchester, and is published with their permission. The writer wishes to thank Mrs. K. V. H. Booth for communicating to him a preliminary version of her manuscript before publication.

REFERENCES

(1) SKUTSCH, R. *Ann. Phys. Chem.*, **61**, p. 190 (1897).
(2) BOOTH, A. D., and BOOTH, K. V. H. *J. Appl. Phys.* To be published.
(3) RAYLEIGH, LORD. *Theory of Sound*, 2nd ed., Chapter VI (London: Macmillan and Co., Ltd,, 1926).
(4) LAMB, H. *Dynamics*, 2nd ed., §50 (London: Cambridge University Press, 1946).

A note on the selection of observers for primary standard optical pyrometry

By M. C. Probine, M.Sc., and Mrs. S. Bertaud, B.Sc., Dominion Physical Laboratory, Department of Scientific and Industrial Research, New Zealand

[*Paper received* 16 *November*, 1953]

The effect of the spectral sensitivity of the eye of the observer on the accuracy with which the International Temperature Scale can be realized above the gold point, by the present instrumental method, is discussed.

The International Temperature Scale is defined above the gold point by the formula[1]

$$\frac{J_t}{J_{Au}} = \frac{\exp\left[C_2/\lambda(t_{Au} + T_0)\right] - 1}{\exp\left[C_2/\lambda(t + T_0)\right] - 1} \qquad (1)$$

where

J_t, J_{Au} = radiant energies per unit wavelength interval at wavelength λ cm, emitted per unit time by unit area of a black body at the temperature t °C, and at the gold point t_{Au} °C, respectively;

$C_2 = 1{\cdot}438$ cm °K;

T_0 = temperature of the ice point in °K;

λ = a wavelength of the visible spectrum.

The above formula holds only for monochromatic radiation. In practice, the scale is extended above the gold point by means of a primary-standard disappearing-filament optical pyrometer. The radiation used covers a band of wavelengths which depends on the transmission of the red filter in the pyrometer, the spectral sensitivity of the observer's eye and the spectral energy distribution of the source. As the radiation is not monochromatic a value of λ for use in the above formula must be defined.

The difficulty has been overcome by defining the "effective wavelength," λ_e, of the pyrometer, for any two temperatures, as the wavelength such that the ratio of the integral luminosities of the black bodies, viewed through the filter, at these temperatures is equal to the ratio of the radiation intensities per unit wavelength interval at wavelength λ_e, i.e.

$$\frac{J_2}{J_1} = \frac{L_2}{L_1} = \frac{\int_0^\infty J(\lambda)_{T_2} \,.\, R(\lambda) \,.\, V(\lambda) \,.\, d\lambda}{\int_0^\infty J(\lambda)_{T_1} \,.\, R(\lambda) \,.\, V(\lambda) \,.\, d\lambda} \qquad (2)$$

where

J_1, J_2 = radiant energies per unit wavelength interval at wavelength λ, emitted per unit time by unit area of a black body at the temperatures T_1, T_2 respectively;

L_1, L_2 = integral luminosities of black bodies at temperatures T_1, T_2 respectively, when viewed through the pyrometer red filter;

$J(\lambda)_{T_1}$, $J(\lambda)_{T_2}$ = spectral distributions of a black body at T_1, T_2 respectively;

$R(\lambda)$ = spectral transmission of the red filter;

$V(\lambda)$ = relative luminous efficiency of a monochromatic radiation.

Having found L_1 and L_2, λ_e for the temperature interval T_1 to T_2 can be calculated from

$$\frac{L_2}{L_1} = \frac{\exp\left[C_2/\lambda_e T_1\right] - 1}{\exp\left[C_2/\lambda_e T_2\right] - 1} \qquad (3)$$

Only approximate values for T_1 and T_2 are required when using the above formulae. In equation (2) the functions $J(\lambda)_{T_1}$ and $J(\lambda)_{T_2}$ can be calculated from Planck's law, and $R(\lambda)$ is obtained by measurement of the filter-glass transmission.

A difficulty arises in the choice of the function $V(\lambda)$. The response of the human eye to radiation of different wavelengths varies from individual to individual and also with intensity, condition of stimulation and, possibly, the vitamin A content of the body. These variables can be eliminated to some extent by the use of pure cone vision (i.e. by limiting the angular aperture to less than 2°, since at the fovea the spectral sensitivity curve is reasonably constant for a wide range of high illuminations). Ideally the function $V(\lambda)$ should be that for the eye of the observer at the time of measurement, and for the condition of measurement. In practice it is usual to use the C.I.E. (1924) curve for the relative luminous efficiency of a monochromatic radiation, which is derived from data recorded by several investigators using different large groups of observers. This curve applies more particularly to foveal photopic vision and therefore, approximately, to the conditions of optical pyrometry. However, in view of the variations among groups of "normal" observers, it seemed worth while to investigate the effect which departures from the standard luminosity curve would have on the "effective wavelength," and therefore on the reproducibility of the scale above the gold point.

CALCULATIONS AND RESULTS

The eye sensitivity curves for each of fifty-two observers have been published by Gibson and Tyndall,[2] on whose work the C.I.E. standard luminosity curve was based. For thirty-eight of these observers the values of the function $V(\lambda)$ are given at each $0{\cdot}01\ \mu$ from $0{\cdot}43$ to $0{\cdot}74\ \mu$. These figures have been used to compute the difference between the platinum-point temperature calculated from the C.I.E. visibility function $V(\lambda)$ and the temperature calculated from the individual visibility functions (hereafter referred to as ΔT_{pt}) of the thirty-eight observers. The values of the function $V(\lambda)$ were published for each $0{\cdot}01\ \mu$, and for our calculations the intermediate values at each $0{\cdot}005\ \mu$ were obtained by graphical interpolation and extrapolation. The figures for the transmission of the filter published by Fairchild, Hoover, and Peters[3] have been used throughout. The values of the

functions $J(\lambda)$, $R(\lambda)$, and $V(\lambda)$ were tabulated from 0·61 to 0·755 μ at each 0·005 μ, and the integrals

$$\int_{0\cdot61}^{0\cdot755} J_1(\lambda) \,.\, R(\lambda) \,.\, V(\lambda) \,.\, d\lambda \text{ and } \int_{0\cdot61}^{0\cdot755} J_2(\lambda) \,.\, R(\lambda) \,.\, V(\lambda) \,.\, d\lambda$$

obtained by using Simson's rule. The calculations were repeated for three observers, using Weddle's integration formula, and no significant difference was obtained.

The calculated values for ΔT_{pt} lay between $\pm 1\cdot5^\circ$ C (with three exceptions).

The significance of this quantity may be assessed by comparing it with the other uncertainties in the measurement. The following table[4,5] gives the estimated errors in two determinations of the platinum point.

Summary of estimated errors

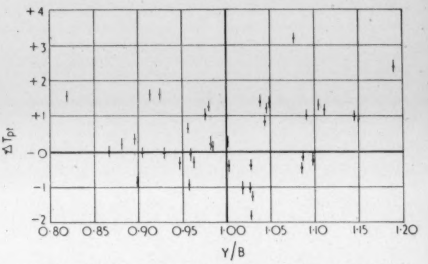
Source of error	Roeser, Caldwell and Wensel[4] Sector No. 2	Sector No. 3	Schofield[5]
Transmission of sector	0·1° C	0·4° C	0·4° C
Effective wavelength	0·4*	0·4	0·4†
Photometric matching	0·5	0·5	0·6
Temperature gradients and lack of black body conditions	0·3	0·3	
Impurity of metal	0·2	0·2	
Maximum if all of one sign	1·5	1·8	1·4

* "The differences between the values for individual observers point to errors in the effective wavelengths used which result in differences from the mean of 0·5°, 0·3° and 0·8° C respectively. The final average, however, should not be in error from this cause by more than 0·4° C."[4]

† "The allowance for effective wavelength is intended to represent the probable error in the transmission curve for the glass. It has been assumed that the visibility curves for two observers are identical. It is hardly practicable to determine the visibility data for each individual engaged in work of this type, but the occurrence of any abnormality in the present work is rendered unlikely by the fact that the Y/B ratios of the two observers were found to be 1·00 and 1·02 respectively and that no discrepancies appeared in their readings when working with or without sectors."[5]

Because the Y/B ratio was taken as a criterion of normality for this purpose, this ratio was calculated for each observer using Gibson and Tyndall's[2] values of the function $V(\lambda)$, and the figures published by Gibson[6] for the spectral transmission of the Y and B test solutions. In the figure ΔT_{pt} has been plotted against the Y/B ratio for each of the thirty-eight observers. The tolerance on ΔT_{pt} is intended to represent uncertainties in the numerical integration. There appears to be no correlation between the Y/B ratio and ΔT_{pt}.

It is also interesting to consider the case of an observer with seriously defective vision. The temperature value at the platinum point has been computed using the eye sensitivity curve of a protanomalous observer (observer D) as determined by McKeon and Wright.[7] The figures given by McKeon and Wright extended towards the red only as far as 0·65 μ. For the purposes of our calculations their results have been extrapolated as far as 0·725 μ by plotting on semi-logarithmic paper and using a straight-line extrapolation.



ΔT_{pt} for thirty-eight observers plotted against Y/B

The curve is therefore an arbitrary one, but it is based on known values for a real observer over part of the curve. The temperature value at the platinum point computed for this observer differed from the C.I.E. value by 5·1° C.

CONCLUSIONS

The calculations of ΔT_{pt} do not reveal the possibility of any serious systematic error in previous determinations of the platinum point, but they do indicate that the accuracy of the present instrumental method of extending the scale above the gold point cannot be improved without special attention to the selection of observers for this work.

ACKNOWLEDGEMENTS

The work described in this paper was carried out in the Dominion Physical Laboratory of the New Zealand Department of Scientific and Industrial Research, and is published with the permission of the Secretary of the Department.

REFERENCES

(1) Hall, J. A., and Barber, C. R. *Brit. J. Appl. Phys.*, **1**, pp. 81–86 (1950).
(2) Gibson, K. S., and Tyndall, E. P. T. *Nat. Bur. Stand. Sci. Pap.*, **2**, pp. 131–191 (1923).
(3) Fairchild, C. O., Hoover, W. H., and Peters, M. F. *J. Res. Nat. Bur. Stand.*, **2**, pp. 931–962 (1929).
(4) Roeser, W. F., Caldwell, F. R., and Wensel, H. T. *J. Res. Nat. Bur. Stand.*, **6**, pp. 1119–1129 (1931).
(5) Schofield, F. H. *Proc. Roy. Soc.*, *A*, **146**, pp. 792–817 (1934).
(6) Gibson, K. S. *J. Opt. Soc. Amer.*, **9**, pp. 113–121 (1924).
(7) McKeon, W. M., and Wright, W. D. *Proc. Phys. Soc. [London]*, **52**, pp. 464–479 (1940).

New books

Microwave spectroscopy. By M. W. P. STANDBERG. (London: Methuen and Co.; New York: John Wiley and Sons Ltd.) Pp. vii + 140. Price 9*s.* 6*d.*

The purpose of this book is to make available such information as is required to obtain and to interpret microwave spectroscopic data as far as they refer to rotating molecules. The book is divided into two parts: (i) a theoretical one (86 pages) dealing with the quantum mechanics of rotating molecules, and (ii) an experimental part (26 pages) describing the apparatus to be used. In addition it contains a short but useful bibliography. The main (theoretical) sections assume a fairly extensive familiarity with quantum mechanical techniques, in particular with matrix methods. In fact the text is so brief that the density of mathematical formulae is similar to that in some mathematical tables. The experimental sections, according to the author, assume a fair amount of practical and theoretical knowledge of electrical circuits. In this way the author has succeeded in fulfilling his aim in the frame of a very compact treatment. The book will no doubt be of great use to those who intend to start research in microwave spectroscopy. H. FRÖHLICH

Thermionic valves, their theory and design. By A. H. W. BECK. (London: Cambridge University Press.) Pp. xvi + 570. Price 60*s.*

For some years after the war, all up-to-date books on thermionic devices were of American origin. The situation has now improved somewhat and Mr. Beck's very erudite book plays an important part in improving it further. The book is written in three parts. The first deals with the fundamental physics of electron emission, followed by a short chapter on phosphors. In Part 2 is given the mathematical theory of electronics, covering electrostatic and magnetic fields, behaviour of individual electrons and of electron beams, transit times and fluctuation noise. The final part deals with types of valves, including velocity-modulated tubes, travelling-wave tubes, magnetrons and picture convertors and storage tubes.

Much of the material in such a book is of course available in a less connected form in earlier textbooks, but the reviewer has not seen anywhere so clear and complete an account of the various microwave tubes, particularly those of the travelling-wave type which have become of great importance recently. The author's very great practical experience shows itself in the excellent discussions of such complex phenomena as cathode formation and behaviour, which are usually treated inadequately.

The book is intended for graduate physicists and electrical engineers, and the extensive use of mathematics will be trying even for some of these, though the descriptive parts may be very valuable for considerably less qualified persons.

There are few errors and the book is extremely good value even at its unavoidably high price. J. H. FREMLIN

Low frequency amplification. By N. A. J. VOORHOEVE (Eindhoven: Philips' Technical Library; London: Cleaver Hume Press Ltd.) Pp. xv + 495. Price 50*s.*

In assessing the merits of this book it is well to see for whom the author intended it. It is stated as being aimed at furnishing the practising sound engineer with an insight into the many subjects which he may encounter in his work. The technical and scientific level of the contents are in keeping with these aims. The reader must not, therefore, expect detailed information on every aspect touched upon, enabling him to engage in fundamental design. Rather, the subjects are treated in general terms so that the engineer will know something about them all to give him a better understanding of their principles and function.

A wide field is covered, from the source of signal (microphone or gramophone pick-up) to the loudspeaker and it includes chapters on all the components (such as resistors, capacitors and transformers), on the power supply units and some general information on complete sound systems.

As might be expected, examples and illustrations are confined almost exclusively to Philips' products and apparatus. This is perhaps no great drawback since there is much similarity in many components with those of other manufacturers, but the chapter on amplifier valves does seem to carry rather too many tables and curves of characteristics, and even outline drawings, of the publisher's valves. Each chapter is well supplied with references so that the reader may pursue the various aspects in more detail in original articles. A. J. MADDOCK

Experimental nuclear physics Vol. 2. Editor: E. SEGRÈ. (London: Chapman and Hall Ltd.) Pp. vii + 600. Price 96*s.*

The second volume of this ambitious project adds two treatises on aspects of nuclear physics, for practising nuclear physicists and advanced students, to the five contained in Vol. 1.

Philip Morrison treats his Survey of nuclear reactions from the standpoint of the theory underlying the experimental results. This field is amenable to theoretical treatment, under conditions where the nature of the nucleon itself does not enter the picture, and the article is admirably planned and written. It is a sign of the times that over one quarter is devoted to nuclear reactions at high energy. An excellent Appendix, unfortunately closed in July 1951, gives classified lists of selected references to work on general phenomena and on particular reactions.

The neutron, by Bernard T. Feld, is a much longer work and represents a very successful attempt to survey the whole field of neutron physics, excluding non-nuclear aspects such as effects on macroscopic properties. The main emphasis is on experimental results and methods, including the theory of the methods and with excellent discussion of the theoretical background. There are a few omissions in the treatment of high energy neutrons; for example, some readers of an otherwise excellent chapter on sources may fail to appreciate how powerful the neutrons produced by proton bombardment can be in aiding precise measurements of high energy cross-sections.

These two contributions authoritatively and concisely present a great deal of information which was previously scattered widely in the journals, and together with Vol. 1 they will be enormously valuable to nuclear physicists. A debt of gratitude is owed to the authors, all of whom are active research workers and must have given up a great deal of their time to this work. T. G. PICKAVANCE

Notes and comments

Elections to The Institute of Physics

The following elections have been made by the Board of The Institute of Physics:

Fellows: J. W. Hughes, J. T. Kendall, M. P. Newby, W. K. Sinclair, R. W. Tiffen, S. E. Williams.

Associates: A. Abbott, A. J. Aston, G. R. Booker, H. D. Bush, P. A. Copeland, P. Cowlin, N. T. M. Dennis, P. D. Fochs, J. T. Griffiths, E. O. Hall, A. W. Hanson, J. H. Harris, L. J. Herbst, A. J. Hones, P. A. Iles, P. J. Jutsum, D. J. Kyte, J. Little, J. R. Mallard, A. D. McRonald, W. V. Morgan, F. Morley, J. W. Murphy, D. A. Patient, A. Porter, E. W. Radoslovich, G. D. Riech, A. R. Roberts, B. E. Stern, N. Sullivan, T. Wainwright, J. B. S. Waugh.

Thirty-three Graduates, nine Students and eight Subscribers were also elected.

Report on discrete sources of extra-terrestrial radio noise

We have received a copy of Special Report No. 3 of the International Scientific Radio Union entitled *Discrete sources of extra-terrestrial radio noise*. The study of the discrete sources of extra-terrestrial radio-frequency radiation is of very recent origin and covers a period of only eight years, during which a considerable volume of original research has been done. A clear summary has become necessary and forms the object of the present report. The mechanism by which celestial bodies emit radio waves is still unknown; some of the identification of radio sources with visible objects are still based largely on supposition, while others seem to be more definitely confirmed.

The report, which is published with the financial help of UNESCO, is available from the General Secretariat of the Union Radio-Scientifique Internationale, 42 Rue des Minimes, Brussels (price 11*s.*).

Radiation Research

The new bi-monthly journal entitled *Radiation Research* is the official organ of the Radiation Research Society and uses the term radiation in its broadest sense which specifically includes ionizing, ultra-violet, infra-red, and visible radiations. The first issue, published in February, contains the following contributions: Introductory remarks on the dosimetry of ionizing radiations, by U. Fano; Neutron physics of concern to the biologist, by Burton J. Moyer; X-ray dosimetry; general principles and experimental factors, by Leonidas D. Marinelli; Measurement of slow neutrons and coexisting radiations, by P. S. Harris; Energy requirements for the inactivation of bovine serum albumin by radiation, by Franklin Hutchinson; Molecular product and free radical yields of ionizing radiations in aqueous solutions, by Edwin J. Hart; A theoretical survey of the radiation chemistry of water and aqueous solutions, by Harold A. Dewhust, Aryeh H. Samuel and John L. Magee; The yields of free H and OH in the irradiation of water, by A. O. Allen; Some effects of heavy-particle irradiation of aqueous acetic acid, by Warren M. Garrison, Herman R. Haymond and Boyd M. Weeks; The role of free radicals and oxygen in reactions produced by ionizing radiations, by E. S. Guzman Barron.

The Managing Editor is Dr. Titus C. Evans, College of Medicine, State University of Iowa, Iowa City, Iowa, to whom manuscripts should be submitted. The publishers are Academic Press Inc., 125 East 23 Street, New York 10, New York. The subscription rate is $9.00 per volume (per year).

Ten Year Index of Proceedings of The Institution of Electrical Engineers

The latest *Ten Year Index* (1942–51) issued by The Institution of Electrical Engineers is much larger than previous similar indexes since in 1941 the *Journal*, the forerunner of the *Proceedings*, was divided into three parts, each with its separate annual indexes. More recently there has been a further division to form Part IV of the *Proceedings*.

The new edition not only embraces the regular issues of the four parts but also the special issues, with their separate indexes, of Parts IA, IIA and IIIA, in which the papers presented at various conventions were published.

Like the annual indexes which are its foundation, the *Ten Year Index* has an entry for every author in a group of joint authors, for every speaker in a discussion, and for every key-word in the title of a paper. The *Ten Year Index*, in addition, has an extremely useful feature which would be inappropriate in an annual index: titles of papers are collected together under 40 broad subject-headings, thus forming small bibliographies on all the topics which are of interest to electrical engineers.

The price is £1 5*s.* (post free) and copies are available from the Secretary, The Institution of Electrical Engineers, Savoy Place, London, W.C.2.

Journal of Scientific Instruments

Contents of the June issue

ORIGINAL CONTRIBUTIONS

Papers

A sensitive air manometer. By L. F. G. Simmons.
A device for extending the range of a recording potentiometer. By P. L. Palmer.
A compact console-type electron microscope. By R. S. Page.
An apparatus for measuring small changes in liquid level. By E. Kovačič.
The design of cold neutron filters. By P. A. Egelstaff and R. S. Pease.
An automatic selector switch with low residual e.m.f. By J. Middlehurst and J. K. Braithwaite.
A simple daylight factor meter. By J. Longmore and R. G. Hopkinson.
A current stabilized photomultiplier power supply By P. Fellgett.
A stable d.c. source of low voltage with low internal resistance. By W. R. Beakley.

Laboratory and workshop notes

Electrolytic engraving on glass. By F. Gutmann and L. W. O. Martin.
Adjustable magnetic control for seismographs. By J. J. Dowling.
Current carrying deflector-plates for rotating a beam of charged particles. By P. A. Einstein and A. G. Edwards.
Making small spheres. By R. L. Whitmore.
A note on the hole, slot and plane method. By J. F. Darby.
A precision r.f. switch. By A. M. J. Mitchell.

NOTES AND NEWS

Correspondence

Calibration of infra-red spectrometers in the wavelength region 15–25 μ. From V. Roberts.
A quartz-coated wire Pirani gauge. From J. H. Leck.

Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2*s.* 6*d.* including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

SPECIAL ARTICLE

Diffuse X-ray scattering and the physical properties of crystals*

By W. A. Wooster, Sc.D., F.Inst.P., Department of Mineralogy and Petrology, University of Cambridge

Both the ionization and photographic methods of studying diffuse X-ray reflexions are described. A brief account of the way in which the experimental observations are related to the thermal agitation of the atoms and to the elastic constants is given for the case of cubic crystals. Precipitates of one phase within another often have the geometrical form of platelets or rods. Each of these types of precipitate is associated with a characteristic type of diffuse X-ray reflexion; the former with spikes in reciprocal space and the latter with disks. A review is given of the present state of knowledge concerning the anomalous X-ray scattering from diamond. New observations on the distribution of the spikes over the reciprocal points are given. It can be shown from these results that the centres of the atoms must occupy their correct positions but that the atomic planes parallel to a given cube face do not all scatter X-rays with equal intensity. The surprising conclusion follows from this that some of the carbon atoms have bonds which are not fully accounted for by the usual four tetrahedral bonds.

The diffraction of X-rays by crystals was discovered in 1912 and from that time until 1939 most investigations were concerned with the strong X-ray reflexions which give the characteristic spots of X-ray photographs. Already in 1923 Waller[1] and Faxen[2] in Sweden had published theoretical papers dealing with the effect of atomic vibrations in crystals in producing scattering in directions other than those of the strong Bragg reflexions. Little attention was given to this work until, at the beginning of the last decade, a number of experimental investigations were begun into the nature of certain reflexions which did not obey the usual rules of the strong Bragg reflexions. It was then found that this early theoretical work had correctly attributed this diffuse reflexion to thermal vibrations of the atoms. Further experimental work has shown that the thermally excited vibrations of atoms are not the only cause of diffuse reflexions, but that atomic misfits of various kinds can cause the same kind of reflexion. It now appears that two particular kinds of measurement can be made by observations on the thermally excited vibrations, namely, the measurement of elastic constants[3–6] and the distribution of the frequencies of atomic vibration at any given temperature.[4,7] The latter type of observation can only be performed with elements or compounds of very similar elements.[8] The study of atomic misfits can be applied to precipitates such as occur in age-hardening alloys,[9] to certain order–disorder problems[10] and, in one case at least, to a study of the nature of chemical binding. Illustrations of some of these uses of diffuse reflexions will be given here.

THE DISTURBANCE OF A LATTICE BY ELASTIC WAVES AND THE CORRESPONDING DIFFRACTION SPECTRA

There is a close analogy between the scattering of visible light from ordinary ruled gratings and the scattering of X-rays from regularly arranged atoms. It is therefore convenient to discuss, firstly, the diffraction of visible light from gratings which have been perturbed by imposing wave-motions on them. A ruled grating consisting of parallel lines gives a series of spectra of the well-known kind on either side of the undeflected beam. To obtain the spectra shown in the following figures the gratings were drawn on cards, photographically reduced and a dozen or so copies were mounted side by side in register with one another. This mosaic was again photographically reduced and the resulting grating was placed next to the lens in a simple apparatus[11] consisting of a lamp, pinhole, grating, lens of long focus and a camera. If a pinhole is used as a source of light, the diffracted beams are restricted to a plane perpendicular to the lines of the grating and passing through the undeflected beam [Fig. 1]. If we impose on the grating a disturbance due to a simple harmonic wave travelling in a direction inclined to the lines of the grating, the lines become wavy [Fig. 2(*a*)]. The wavefronts are shown in Fig. 2(*b*) and along any one front the points on the lines of the grating are all displaced by the same amount in the direction of the wave-normal. The spectra obtained from such a grating [Fig. 2(*c*)] using a pinhole source, are three times as numerous as those obtained with the straight line grating. Each previous spectrum is flanked by two others, one on each side, and the line through the groups of three is perpendicular to the wavefronts. The longer the wavelength of the waves the nearer are the extra spectra to the set corresponding to an undisturbed straight line grating.

A two-dimensional grating, which more closely resembles the distribution of atoms in crystals, can be constructed from a series of dots arranged on a lattice [Fig. 3(*a*)] and the spectra given by this array [Fig. 3(*b*)] is similar, but with the long edge of the repeat unit parallel to the short edge of the repeat unit in the grating.

We can now suppose waves to be travelling across such an array of dots, first in a direction parallel to one edge of a repeat unit. If the individual dots move in the direction of the vertical wave-normal [Fig. 4(*b*)] the waves are longitudinal. Such a perturbed lattice [Fig. 4(*a*)], gives spectra similar to those of the unperturbed lattice except that each spectrum is now split into three [Fig. 4(*c*)]. The line passing through each group is perpendicular to the waves. The spectra lying on a horizontal line through the centre, parallel to the waves, are not affected by the waves. The reason they are not split is that the spots are displaced along the diffracting planes and these planes are therefore quite flat and of the same spacing as in the unperturbed lattice.

If the waves travel in a direction inclined to the side of the repeat unit [Fig. 5(*b*)], each horizontal and each vertical row of dots becomes wavy [Fig. 5(*a*)]. The spectra which result from such a perturbed grating are again split into three [Fig. 5(*c*)], except when the displacement of the points is perpendicular to the line joining the corresponding spectrum to the undeflected beam.

* Based on a lecture given on 29 May, 1953, before the X-ray Analysis Group at The Institute of Physics Convention in Bournemouth.

To sum up, a sinusoidal wave superimposed on a two-dimensional grating produces two extra spectra flanking each original spectrum and the line passing through the three spectra is parallel to the wave-normal. The distance of the secondary maxima from the primary is inversely proportional to the length of the perturbing wave. The only exception to this arises when the particle displacement is actually in the planes giving a particular spectrum.

THE THERMAL DIFFUSE REFLEXION OF X-RAYS

In order that the diffuse reflexions can be observed conveniently, the elastic waves should have wavelengths lying between ten and one hundred times the repeat distance of the crystal lattice. The cell size of most crystals is between 5 and 15 Å and thus the range of wavelengths involved in our discussion is about 50 to 400 Å. The range can be increased both up and down by greater attention to the details of the observation. The frequency of such waves is of the order of 10^{11} c/s and this is far higher than can be generated by any piezoelectric or other mechanical device. Fortunately, however, the normal movement of the atoms—their thermal agitation—at all temperatures above absolute zero provides waves of just this range of frequency. The vibration of atoms in crystals is usually thought of as being a random motion and from the point of view of any one atom that view is correct. From the point of view of the whole lattice, however, the atomic movements can be regarded as the resultant of a large number of waves travelling in all directions and with a great range of wavelengths and frequencies.

It is not a matter of chance how large is the amplitude of any particular wave. The amplitudes are regulated by the all-embracing laws of thermodynamics; Planck's fundamental postulate, which has been confirmed in so many branches of science, requires that, at ordinary and higher temperatures, each mode of vibration, in a series of modes all in equilibrium with one another, shall have an energy equal to kT. The energy of a wave-motion is also given by $E = \frac{1}{2}a^2n^2$, where a is the amplitude and n the frequency. Thus for any particular mode of vibration we have

$$\tfrac{1}{2}a^2n^2 = kT. \tag{1}$$

Now any wave motion also has a relation,

$$n\Lambda = C \tag{2}$$

between frequency n, wavelengths Λ and velocity C.

Thus the amplitude of the wave is related to its velocity by the combination of equations (1) and (2):

$$a^2 = 2kT/n^2 = 2kT(\Lambda^2/C^2) \tag{3}$$

It is obvious on general grounds that, for small amplitudes, the greater the amplitude of the wave perturbing the lattice the greater will be the corresponding intensity of the scattered X-rays. In fact, the intensity of the diffuse beam is proportional to a^2. Thus from the measurement of intensity of the diffusely scattered beam we obtain Λ^2/C^2. Now Λ is fixed by the direction in which the scattered X-ray beam is travelling. In the optical spectra shown earlier, the distance of the extra spectra from the nearest primary spectrum was inversely proportional to the wavelength of the perturbing waves, and the extra spectrum occurred on a line through the primary spectrum parallel to the normal to the perturbing waves. Similarly, by observing the X-rays travelling at a known angle to a given normal reflexion, we can fix both the direction and the wavelength of the elastic waves responsible for that particular diffuse beam. We shall return to this point later in connexion with the experimental arrangements. Thus we have found Λ^2/C^2 from the intensity and Λ from the direction of the observation and hence we obtain the velocity C.

The relation between the velocity of an elastic wave travelling through a crystal and the elastic constants was worked out long ago. Nearly eighty years ago[12] the general treatment was developed which related, for any direction of travel of an elastic wave (*a*) the velocity, (*b*) the direction of vibration, and (*c*) the elastic constants of the crystal. Except along certain directions, the velocity is dependent upon several elastic constants. However, the velocity of longitudinal waves travelling along, say, the [100] axis of a cubic crystal, depends only on one elastic constant, namely, c_{11}. The relation between the velocity of this wave and c_{11} is given by the equation

$$C = \sqrt{(c_{11}/\rho)} \tag{4}$$

where ρ is the density. As we have seen in the optical diffraction due to longitudinal waves travelling parallel to one edge of the unit cell [Fig. 4] the extra spectra lie on the line through the normal spectrum in the direction of the wave-normal. Similarly, with X-ray diffraction, the diffuse scattering corresponding to the longitudinal waves travelling along [100] arises from points in reciprocal space which lie near the reciprocal point 100 on the line joining this point to the origin [Fig. 6(*a*)]. Transverse waves having the wave-normal along [010] and the vibration direction along [100] have a velocity also determined by one elastic constant, namely, c_{44} according to the equation,

$$C = \sqrt{(c_{44}/\rho)} \tag{5}$$

Such waves produce diffuse reflexions in directions corresponding to reciprocal points which lie on the line through the reciprocal point 100, parallel to [010], which is the direction of the wave-normal [Fig. 6(*b*)]. By setting the crystal so as to observe these two diffuse reflexions the values of c_{11} and c_{44} may be independently determined for a cubic crystal. The only remaining elastic constant is c_{12}, and this may be found, in combination with c_{11} and c_{44}, by measuring the diffuse reflexion at a reciprocal region near to the point 100 lying on a line parallel to [110] or to [111] passing through the point 100.

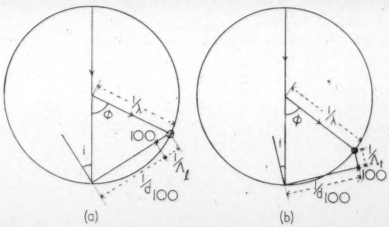
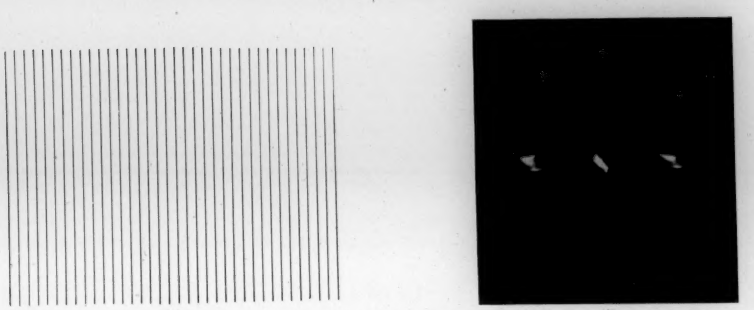


Fig. 6. (*a*) The small shaded circle represents the region of reciprocal space which lies on the intersection of the vector joining the origin to the point 100 and the reflecting circle. The setting for reflexion by longitudinal waves. (*b*) The setting required for reflexion by transverse waves

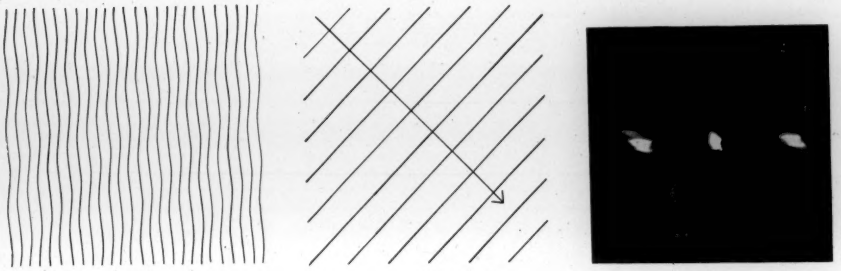
EXPERIMENTAL METHODS

(i) *Ionization- and counter-spectrometer methods.* Just as in the early study of Bragg reflexions, the quantitative results



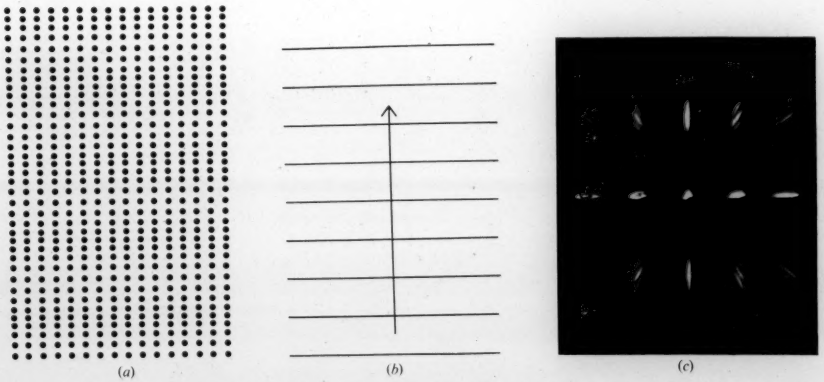
(*a*) (*b*)

Fig. 1. (*a*) Line-grating from which the spectra shown in (*b*) were obtained



(*a*) (*b*) (*c*)

Fig. 2. (*a*) Line-grating perturbed by waves shown in (*b*), supposed moving in the direction of the arrow. (*c*) Main and subsidiary spectra produced by this grating



(*a*) (*b*) (*c*)

Fig. 4. (*a*) Grating of dots perturbed by longitudinal waves shown in (*b*). (*c*) Main and subsidiary spectra produced by grating

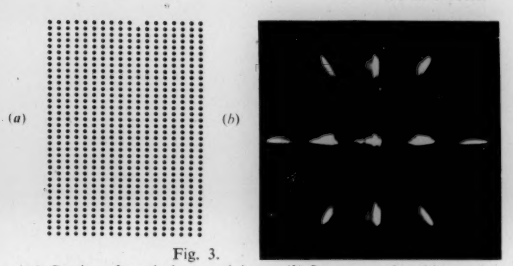


Fig. 3.
(*a*) Grating of regularly spaced dots. (*b*) Spectra produced by grating

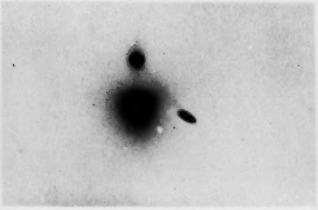


Fig. 8. Diffuse reflexion photograph of lithium fluoride

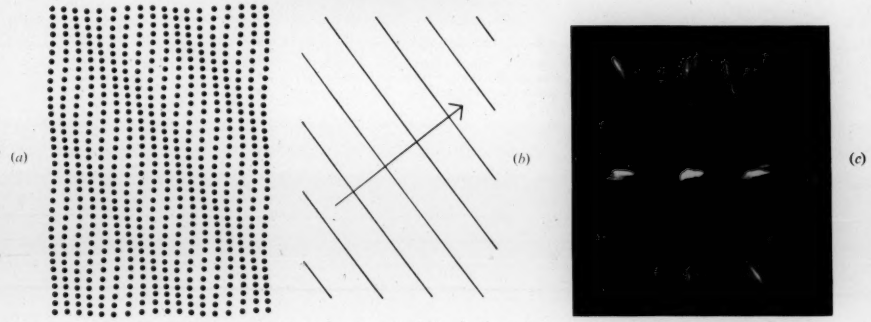
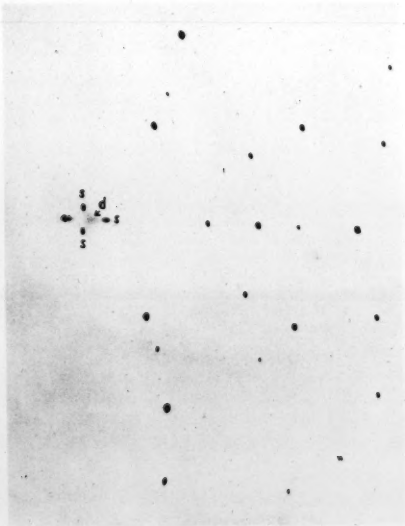


Fig. 5. (*a*) Grating of dots perturbed by waves travelling diagonally as shown in (*b*). (*c*) Main and subsidiary spectra produced by grating



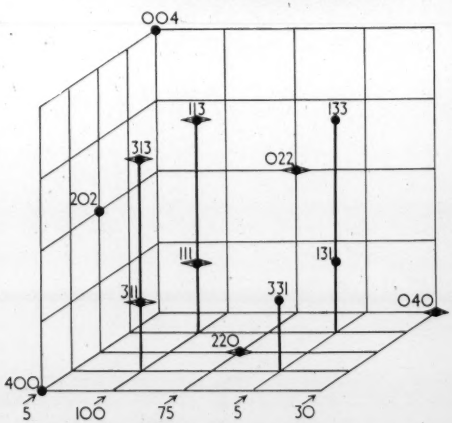
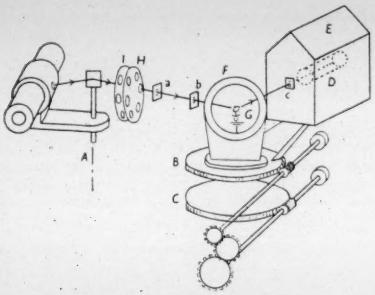


Fig. 11. The spikes parallel to [010] associated with the reciprocal lattice points of diamond

Left.—Fig. 10. Laue photograph of diamond showing spike-spots, *s*, and thermal diffuse spot, *d*, associated with the 111 reflexion

on diffuse reflexions were first obtained by an ionization spectrometer and then, later, photographic methods were evolved to give the same information.[6,13] Fig. 7 shows in diagrammatic form the parts of the spectrometer used in Cambridge—X-ray tube, monochromatizing curved crystal, absorbing screens, *I*,*H*, collimating slits, *a*,*b*, a vertical circle supporting the goniometer head *G* and finally an ionization chamber, Geiger counter, or proportional counter, *D*. The



[Reproduced from *Acta Crystallographica*]

Fig. 7. Component parts of the Geiger-counter spectrometer; *A*, axis about which X-ray tube and monochromatizing crystal pivot; *B*, *C*, worm wheels supporting counter-box, *E*, and vertical circle, *F*; *D*, Geiger–Müller tube; *G*, goniometer head; *H*, *I*, disks supporting absorption screens, *a*, *b*, *c*, limiting slits

movements of the crystal and the counter are controlled by gear wheels which enable the setting to be accurately made. All reflexions which are recorded occur in a horizontal plane, and the settings of the crystal and the counter determine the point in reciprocal space which is under observation.

Apart from the geometrical requirements of the experiment there is a rather difficult problem concerned with the measurement of weak X-ray intensities. In order to make the results unambiguous it is best to reflect the incident radiation from monochromatizing bent quartz crystals. The intensity of the diffuse radiation coming from the crystal under test is usually weak, amounting to perhaps not more than thirty quanta per minute. By careful attention to the design of electronic circuits and counters it is possible to reduce the background counting rate to about eight per minute and so the measurements become possible. From what has been said earlier, it will be realized that to determine the amplitude of the elastic waves it is necessary to know what fraction of the incident X-ray intensity is contained in the diffusely scattered beams. Now the ratio of these two intensities is of the order of 10^6 and it is consequently difficult to make this comparison with accuracy. In certain instances the problem is simplified by the presence of Compton scattering from the crystal, of an intensity comparable with the diffuse intensity. The Compton scattering varies little with the setting of the crystal, whereas the diffuse scattering due to thermal waves varies rapidly as we depart from a reciprocal point. It is therefore easy to determine the strength of the Compton radiation. When this has an easily calculable value, as, for instance, in diamond, it can be used to standardize the intensity of the direct beam.

(ii) *Photographic method.* Diffuse reflexions are usually studied photographically by means of a stationary crystal and a stationary film, using either filtered white radiation or crystal-reflected monochromatic radiation. The crystal is set so that a given plane is near the setting required for a Bragg reflexion. If filtered radiation is used, a Laue photograph is obtained and one or more diffuse spots are obtained at or near the places on a film where the Bragg reflexion would have occurred if the crystal had been set to give that reflexion. The amount of the mis-setting is usually between 0°·5 and 4°·0. It may be convenient to mount the crystal on a Weissenberg goniometer so that after the diffuse spot has been recorded the crystal can be rotated so as to give the Bragg reflexion slightly to one side of the diffuse spot. Such a photograph obtained with lithium fluoride is shown in Fig. 8. By simple linear measurement on the diffuse spot it is possible to identify the points which correspond to various directions in reciprocal space. These points within the diffuse spot are studied photometrically and the corresponding intensities of the diffuse reflexions determined. Thus from the position of the point within the diffuse spot we determine the direction of the wave-normal and the wavelength of the elastic wave responsible for the diffuse reflexion. From the density at the points considered the velocity can be determined as already explained and finally, from this, one only, or a combination of elastic constants can be found.

THE STUDY BY DIFFUSE REFLEXIONS OF PRECIPITATES IN SOLID CRYSTALS

Diffuse reflexions may be caused by perturbations of the regular lattice due to causes other than the elastic waves. If, for instance, a regular cubic lattice is distorted by the inclusion of a thin platelet of foreign atoms, slightly larger than those of the matrix [Fig. 9(*a*)], then a diffuse scattering is obtained which corresponds in reciprocal space to spikes passing through each normal reciprocal lattice point (relp) [Fig. 9(*b*)]. These spikes are all directed perpendicular to the plane of the platelet and decrease in intensity outwards from

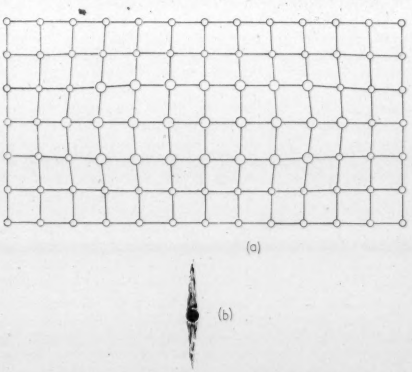


Fig. 9. (*a*) Square lattice distorted by a platelet composed of larger atoms than those of the matrix. The plane of the platelet is perpendicular to that of the diagram. (*b*) Normal reciprocal point with spikes radiating from it along a line normal to the plane of the platelet

the relp. That such a series of spikes is to be expected can be seen in the following way. The distortion of the lattice can be resolved into the displacements due to a series of wave-trains of differing wavelengths but all having their wave-normals perpendicular to the plane of the platelet. Each of these wave-trains produces a pair of scattering points in reciprocal space at equal distances on either side of the relp, and lying on the line through the relp perpendicular to the plane of the platelet. All of these diffuse scattering points together form the continuous spikes. Such spikes are found in many alloys, particularly the age-hardening alloys, e.g. 4% copper in aluminium. If the precipitate takes the form of a rod distorting the normal matrix, the normal relps have diffuse disks around them and the plane of the disks is perpendicular to the length of the rod-like precipitate. The distortion of the atomic arrangement can be resolved into a series of wave-trains all having wave-normals perpendicular to the length of the precipitate. The diffuse scattering points corresponding to these waves must be restricted to a plane as already described.

The physical properties of many alloys are profoundly affected by the deposition of such disks and rods. One of the most important justifications for the study of diffuse reflexions is that they may be registered before the effect of precipitation of impurities can be detected either microscopically or by a change in the intensity of Bragg reflexions. The early stages of precipitation which can best be followed by diffuse reflexions are associated with the development in various alloys of greater hardness and tensile strength and in other, magnetic, alloys with greater remanence.

CHEMICAL BINDING AND DIFFUSE REFLEXION IN DIAMOND

In recent years it has been shown that although diamonds differ among themselves, the majority give rise to two types of diffuse reflexion as is shown in Fig. 10.[14-16] This is a Laue photograph and most of the spots are quite usual Bragg reflexions. However, the spots near to the 111 reflexion are abnormal and it can be seen that they are of two types, namely, sharp and diffuse, denoted in Fig. 10 by *s* and *d* respectively. The diffuse spot *d* is due to the thermal vibrations and from its study the elastic constants have been found by the method already described.[17] The sharp spots correspond to the spikes passing through the relp 111 parallel to the directions [100], [010], [001]. Many investigators have studied these spikes and the facts now available can be summarized by Fig. 11.[18] Although the spikes passing through any relp may be parallel to one, two or three of the directions [100], they are only drawn, in Fig. 11, parallel to [010]. This is intended to simplify the presentation and probably corresponds to the actual state of affairs in a sufficiently small portion of the crystal. The length of the spikes in Fig. 11 indicates the ratio of the spike intensity to that of the corresponding relp. It will be seen that this ratio is constant in any reciprocal plane parallel to (010), but that its value varies with the distance of the reciprocal plane from the origin. In the table the spike intensity is given as a function of the index *k* of the reciprocal plane in which the relp lies.

Table showing the spike intensity in successive 0*k*0 *planes*

Index k	0	1	2	3	4	5
Spike intensity	5	100	75	5	30	0

The total number of spikes present in any actual crystal may be obtained by applying the same rules to the directions [100] and [001] in turn, since the crystal possesses a three-fold axis of symmetry parallel to [111].

A TENTATIVE THEORY OF THE SPIKES IN DIAMOND

A number of research workers have advanced theories designed to explain the diffuse reflexions of spike-like character in reciprocal space given by diamond.[19-21] None of these theories is wholly satisfactory and the theory put forward here must be regarded as tentative because it is somewhat revolutionary in character.

This theory is based on two principal assumptions, the former originally advanced by Guinier[20]:

(i) that the whole crystal is composed of "domains" within which the spikes are arranged parallel to only one of the three directions [100] as is illustrated by Fig. 12. These domains are larger than about 10^{-4} cm in diameter and smaller than about 10^{-2} cm.

(ii) the carbon atoms do not all possess four bonds which are exactly equivalent in their scattering power for X-rays. This is the essentially revolutionary assumption.

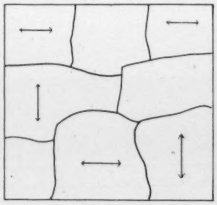


Fig. 12. Illustration of the proposed domain texture of diamond

The direction of the spikes associated with any given domain will be called the "unique direction of the domain". If, the unique direction is [100], then the carbon atoms in any one atomic plane parallel to (100) are assumed to be all alike in their power of scattering X-rays, but slightly different in scattering power from other parallel planes of carbon atoms. This difference in the scattering power, if randomly distributed through the domain, must produce the observed diffuse reflexion corresponding to the spike. An optical illustration of this phenomenon can be constructed as shown in Fig. 13. On the left-hand side, the diagram represents the atomic lattice of diamond projected on to a (100) plane. The optical diffraction pattern produced in the apparatus described earlier by this grating is shown below it. To imitate the variation of scattering power, vertical white lines have been drawn across the rows of dots. The distribution of these white lines was random. The diffraction pattern produced by this arrangement is shown below it and the optical analogues of the spikes, perpendicular to the vertical white lines, can be seen.

Chemists and crystallographers have, until now, assumed that the four bonds joining any one carbon atom in diamond to its neighbours, are equivalent. The new feature of the theory put forward here is that they are not quite equivalent. There might be two or even three kinds of bonds between neighbours and the differences in scattering power between the planes of atoms flanked by the different kinds of bonds would be small. Such a random distribution of non-equivalent planes of atoms as that assumed here provides a qualitative explanation of the observed spikes.

Quantitative measurements have also been made to determine (i) the rate of variation of the intensity of the spike along its length and (ii) the ratio of the spike intensity to that

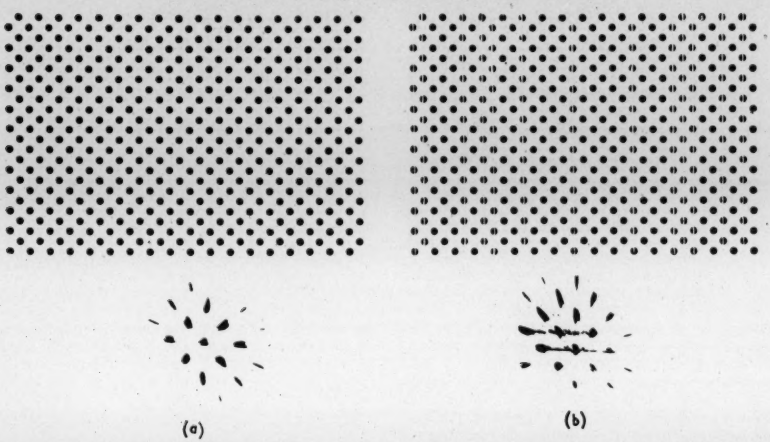


Fig. 13. Optical analogue of the spikes in diamond: (*a*) (*above*) lattice with all planes parallel to (010) of equal scattering power; (*below*) corresponding optical diffraction pattern; (*b*) (*above*) similar lattice but with planes parallel to (010) of unequal scattering power and randomly distributed; (*below*) corresponding optical diffraction pattern showing spikes

of the corresponding Bragg reflexion. The results of the first measurement show that if R is the distance in reciprocal space of a point on the spike from its relp and I is the intensity of the spike at this point, then

$$I \propto \frac{1}{R^{2\cdot 2}} \qquad (6)$$

The second measurement gives a result which varies greatly from one crystal to another, but a strong spike would correspond to a ratio of about 1 : 1500.

If we apply Wilson's theory[22] to the model of the atomic arrangement described above we have the relation

$$I/NJ_0 = 2\zeta/[\zeta^2 + (2\pi\omega)^2] \qquad (7)$$

in which NJ_0 is the intensity of the Bragg reflexion, ζ is the probability that a given plane of atoms perpendicular to the unique axis of the domain will be irregular and ω is the distance along the spike from the relp expressed as a fraction of the length of the reciprocal unit cell. Using the value $I/NJ_0 = 1/1500$ and $\omega = 0{\cdot}1$, we obtain $\zeta = 2{\cdot}6 \times 10^{-4}$. Thus the probability of an irregular arrangement of the bonds is about 1 in 4000. On this theory the exponent of ω should be the same as that of R, whereas the experimental limits to the exponent of R are $2{\cdot}1 - 2{\cdot}3$, and do not include the theoretical value $2{\cdot}0$. This discrepancy suggests that the assumption of a random distribution of planes of abnormal scattering power is not quite justified.

In conclusion, it should be said that several other types of measurement can be made using diffuse X-ray reflexions. The most important of these are (i) the determination of the distribution of frequencies of waves in elements and alloys of similar elements and (ii) the study of order-disorder problems. The first measurement is a direct extension of the work described under the study of elastic constants and the second is closely allied to the study of precipitates. It is perhaps surprising that so many properties of crystals can be studied by a single technique based on such weak X-ray reflexions.

REFERENCES

(1) Waller, I. *Z. Phys.*, **17**, p. 398 (1923); *Z. Phys.*, **51**, p. 213 (1928).
(2) Faxen, H. *Z. Phys.*, **17**, p. 266 (1923).
(3) Laval, J. *Thesis* (Paris, Sorbonne, 1939).
(4) Olmer, Ph. *Bull. Soc. Franç. Min.*, **71**, p. 145 (1948).
(5) Ramachandran, G. N., and Wooster, W. A. *Acta Cryst.*, **4**, pp. 335, 431 (1951).
(6) Lonsdale, K., and Smith, H. *Nature* [*London*], **149**, p. 21 (1942).
(7) Curien, H. *Acta Cryst.*, **5**, p. 393 (1952).
(8) Cole, H., and Warren, B. E. *J. Appl. Phys.*, **23**, p. 335 (1952).
(9) Guinier, A. *Acta Cryst.*, **5**, p. 121 (1952).
(10) Warren, B. E., Averbach, B. L., and Roberts, B. W. *J. Appl. Phys.*, **22**, p. 1493 (1951).
(11) Bragg, W. L. *Nature* [*London*], **154**, p. 69 (1944).
(12) Christoffel, E. B. *Ann. di. Mat.*, **8**, p. 193 (1877).
(13) Hoerni, J., and Wooster, W. A. *Acta Cryst.*, **5**, p. 626 (1952); **6**, p. 543 (1953).
(14) Lonsdale, K. *Proc. Roy. Soc. A*, **179**, p. 315 (1942).
(15) Raman, C. V. *Curr. Sci.*, **12**, p. 33 (1943).
(16) Krishnan, R. S., and Ramachandran, G. N. *Nature* [*London*], **155**, p. 234 (1945).
(17) Prince, E., and Wooster, W. A. *Acta Cryst.*, **6**, p. 450 (1953).
(18) Hoerni, J., and Wooster, W. A. *Experientia*, **8**, p. 297 (1952).
(19) Raman, C. V., and Nilakantan, P. *Proc. Indian Acad. Sci. A*, **14**, p. 317 (1941).
(20) Guinier, A. *Bull. Soc. Franç. Min.*, **67**, p. 382 (1944).
(21) Grenville-Wells, H. *Thesis* (University of London, 1951).
(22) Wilson, A. J. C. *X-ray Optics* (London: Methuen and Co. Ltd., 1949).

The use and scope of Iridium 192 for the radiography of steel

By R. HALMSHAW, B.Sc., A.R.C.S., A.Inst.P., Armament Research Establishment, Woolwich, London, S.E.18

[*Paper first received* 23 *September, and in final form* 2 *December*, 1953]

The radiographic techniques and scope of application of ^{192}Iridium for the radiography of steel, are discussed. Sensitivities obtainable have been calculated from the experimentally determined effective absorption coefficient and the ratio of scattered to direct radiation intensities; these sensitivities have been confirmed experimentally. An attempt has been made to assign an equivalent X-ray kilovoltage to the radiation. The definition obtainable and its importance with different types of flaw have been discussed, and from this data the ranges of thickness of welds and castings on which satisfactory flaw-sensitivity should be obtained have been suggested.

The introduction of radio-isotopes has greatly stimulated interest in γ-ray radiography. The simplicity of the apparatus required, together with the relatively low initial outlay compared with X-ray equipment, has led many people to consider the use of γ-rays for work which would normally be done with medium kilovoltage X-rays. It is important, therefore, to examine the suitability of isotopes for different types of radiographic inspection. ^{192}Ir has a γ-ray spectrum of considerably lower energy than ^{60}Co, ^{182}Ta or radon, so that it is likely to be more suitable than these for the examination of thin steel specimens, and this paper is limited to a consideration of this isotope.

QUALITY OF RADIATION FROM ^{192}IR

^{192}Ir is prepared in the atomic pile by a neutron-γ reaction, the target material being metallic iridium. ^{194}Ir is also produced but has a half-life of only nineteen hours. The half-life of ^{192}Ir has been quoted at various figures between 70 and 75 days, but the most reliable value appears to be 74 ($\pm \frac{1}{2}$) days.[1,2] The energy distribution in the γ-ray spectrum is complex, and there appears to be some disagreement on the relative intensities of the spectrum lines.

Table 1 gives the relative intensities as quoted by Morrison[2] and by Müller and others.[3] Some evidence has also been published recently for the existence of a number of spectrum lines at higher energy (0·77, 0·87, and up to 1·2 MeV).[4]

Table 1. *γ-ray spectrum of ^{192}Ir*

Energy of spectrum line (MeV)	*Relative intensity (Morrison)*	*Relative intensity (Müller)*
0·207	45·5	1
0·206		7·5
0·296	91	38
0·308	100	37
0·316		100
0·468	200	30
0·485		1·1
0·588	182	1·1
0·605	55	1·4
0·613		0·5

In order to check these figures, an absorption curve for a heavily collimated beam, determined experimentally, was compared with the curves calculated from the data in Table 1. The calculated curves were obtained by taking an absorption coefficient for each separate spectrum line energy and determining the change in composition of the beam of radiation as the thickness of absorbing material was increased. The resulting curves are shown in Fig. 1, plotted as $\log_{10}$ (exposure, for a constant film density) against thickness; it will be seen

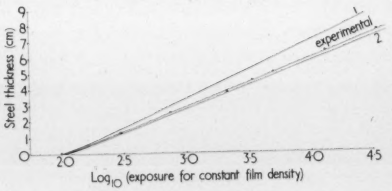


Fig. 1. The absorption in steel of ^{192}Ir radiation under "narrow beam" conditions

Curve 1, calculated from data by Morrison; Curve 2, calculated from data by Müller.

that Müller's data agrees closely with the experimental curve, but that the latter suggests a slightly higher proportion of the higher energy spectrum lines to be present.

EQUIVALENT X-RADIATION

Estimates of the equivalent X-ray voltage are of interest in that they give a rough indication of the range of specimen thickness over which adequate sensitivity can be obtained. By comparing effective absorption coefficients measured from the linear portion of radiographic exposure-curves[5,6] an estimate of the equivalent X-ray radiation can be obtained.

Table 2 shows the effective linear absorption coefficient, measured in this way, taken from various exposure curve data.

Table 2. *Effective linear absorption coefficient*

Radiation	*Effective absorption coefficient (in.$^{-1}$)*
200 kV_p	3·7
400 kV_p	2·2
1 MeV (d.c. circuit)	1·18
2 MeV (d.c. circuit)	0·84
500 kV (d.c. circuit)	1·54
1 MeV (resonance transformer)	1·2
2 MeV (resonance transformer)	0·91
^{192}Ir	1·25

These absorption coefficients are necessarily only approximate, as the effective hardness of the radiation from an X-ray

tube depends to some extent on the electrical characteristics of the generator and the inherent filtration in the X-ray beam, but by plotting μ against 1/kV an approximate straight line is obtained (Fig. 2), and the data indicates that the quality of

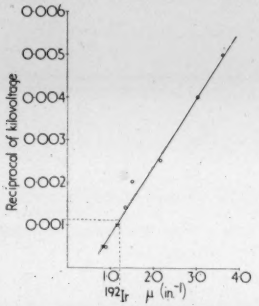


Fig. 2. The relation between effective linear absorption coefficient and kilovoltage

the radiation from ^{192}Ir, after it has passed through an initial thickness of specimen, is roughly equivalent to an X-ray tube running at a peak voltage of 800–900 kV_p.

PENETRAMETER SENSITIVITY

A convenient method of obtaining a measure of the sensitivity of a radiograph is to use a standard type of step-penetrameter. Such penetrameters have sharp straight edges, whereas flaws in the specimen are often of diffuse and irregular shape, but these limitations are generally appreciated and it is considered that the use of this type of penetrameter does permit an assessment of the minimum detectable difference in thickness, and hence, of the potential flaw-sensitivity.

The minimum perceptible brightness-difference which can be perceived by a properly adapted eye, under optimum conditions of brightness, is given by Kruithof[7] as being represented by a density-difference on a film of just under 0·006. This is considerably lower than the figure of 0·02 which is often quoted, but has been confirmed by the author by direct experiment, using a fine-grain emulsion. The experiments were not extended to relatively grainy X-ray film emulsions, but Kruithof[7] has also shown that the visibility of a small density-step is not critically dependent on the sharpness of the density-change across the step. His experiments show that the visibility of the density-change is not affected by the blurring of the edge unless the width of the blurring subtends an angle of more than 8′ of arc at the eye. This value of 0·006, for the minimum perceptible density-difference, has therefore been used in estimating the sensitivities.

An equation for sensitivity can be derived, and is:

$$\text{Sensitivity} = \Delta x/x = -2{\cdot}3\,.\,\Delta D(1 + I_s/I_D)/\gamma_D\,.\,\mu\,.\,x \quad (1)$$

where Δx is the minimum thickness of specimen detectable in a thickness x, ΔD is the minimum discernible density-difference, μ is the absorption coefficient, γ_D is the gradient of the characteristic curve at density D, and I_s/I_D is the ratio of scattered to direct radiation reaching the film. By producing (log E − thickness) curves for conditions of open field (scatter present) and with a collimated beam (direct radiation only), the ratio I_s/I_D can be measured and the value of μ determined.

The above equation includes the gradient of the film characteristic curve, γ_D. The variation of γ_D with density for two typical non-screen type films, one a fast non-screen film and one a slow, fine-grain, non-screen type is shown in Fig. 3. These curves show that the value of γ_D tends to a maximum with the fast type of film so that little is to be gained by exceeding a film density greater than 2·0. On the other hand, with fine-grain films, the value of γ_D continues to

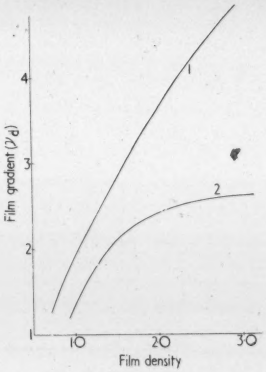


Fig. 3. The variation of film gradient with film density (experimental)

Curve 1, fine grain non-screen type film; curve 2, fast non-screen type film.

increase, with density, up to densities which are too high for effective viewing and a compromise has to be made; in practice a film density of about 3·0 is used. It will be seen, therefore, that fine-grain film has the two-fold advantage of giving better definition because of the fine grain, and of higher sensitivity due to the higher values of γ_D which can be obtained on it.

Sensitivity curves have been drawn, using this equation, with a value of $\Delta D = 0{\cdot}006$. Fig. 4 shows curves for ^{192}Ir radiation, with Kodak Crystallex film, calculated for values of γ_D corresponding to film densities of 1·0, 2·0, 3·0. This film is typical of the fine-grain type and the curves illustrate the rapid fall off in sensitivity for densities below 2·0. Fig. 5 shows sensitivity curves calculated for ^{192}Ir radiation, 200 kV and 400 kV X-rays, based in each case on a value of $\gamma_D = 4{\cdot}0$. 200 kV X-rays would normally be used for the examination of steel thicknesses from ½–1½ in., and 400 kV for 2–3 in., and over these ranges of thickness the sensitivity obtainable with ^{192}Ir is inferior to that with X-rays. For thicknesses of steel greater than 2 in. where 400 kV is more normally used, the sensitivities become almost equal and are in the region of ⅓%, which would be considered very satisfactory for most classes of industrial radiography. The most important fact to be noted from Fig. 5 is, however, the very rapid falling off of sensitivity with ^{192}Ir when used for thicknesses below 1 in. steel. With 1 in. steel ^{192}Ir gives half the sensitivity obtainable with 200 kV and for ½ in. steel this figure drops to two-fifths;

for thickness below ½ in. the rapid falling off in sensitivity with ^{192}Ir becomes increasingly obvious. For such thicknesses a higher sensitivity would generally be obtained by using X-rays of lower energy than 200 kV, so that the loss in sensitivity

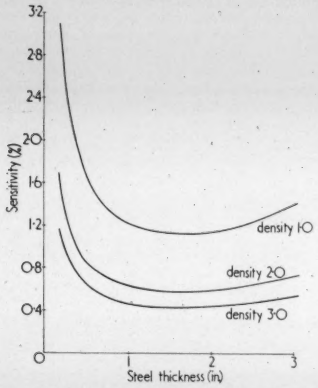


Fig. 4. Penetrameter sensitivity obtainable with ^{192}Ir radiation, against steel thickness, for three film densities

when comparing ^{192}Ir with X-rays is even greater than is shown by Fig. 5. Except for the detection of gross defects the use of ^{192}Ir is not recommended if the specimen thickness is less than ½ in. steel.

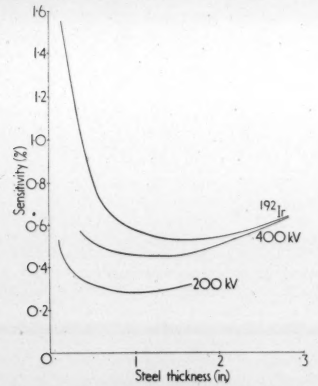


Fig. 5. Penetrameter sensitivity obtainable with ^{192}Ir radiation, 200 kV and 400 kV X-rays, against steel thickness

In order to investigate the accuracy of these sensitivity curves, radiographs were taken of a plate, with grooves varying from 0·008 to 0·072 in. in depth. The plate was built up to different thicknesses and radiographed with the grooves on the side of the specimen remote from the film; films of different densities were produced and the shallowest discernible groove determined. The calculated and observed sensitivities are listed in Table 3.

Table 3. *Sensitivity, as measured with a step-penetrameter*

Radiation	*Steel thickness (in.)*	*Film density*	*Observed sensitivity (%)*	*Calculated sensitivity (%)*
^{192}Ir	0·5	1·05	1·6	1·56
^{192}Ir	1·0	1·5	0·8	0·84
^{192}Ir	1·75	1·2	0·9	1·10
^{192}Ir	2·5	1·7	0·6	0·84
^{192}Ir	1·5	2·7	0·5	0·47
200 kV	1·0	1·5	0·5	0·43
200 kV	1·75	1·2	0·6	0·70
400 kV	2·0	2·1	0·7	0·53

These sensitivities are considerably higher than might be expected, but it must be borne in mind that they are based on the discernibility of long straight regularly-spaced grooves.

INHERENT AND GEOMETRIC UNSHARPNESS EFFECTS

The discussion, so far, has dealt chiefly with the influence of contrast on sensitivity, but the quality of the radiograph, and therefore the flaw-sensitivity obtainable, is also affected by those factors controlling the definition or unsharpness of the image. Definition and contrast are not independent variables affecting the quality of the image, but their combined action is difficult to analyse unless some specific shape or size of flaw is considered. It is proposed, therefore, to consider the question of definition separately.

The dependence of radiographic definition on inherent unsharpness, caused by the scatter of X-rays and electrons within the emulsion, has been discussed by the author in an earlier paper.[8] Measurements of inherent unsharpness have been made with ^{192}Ir radiation and are given in Table 4, together with comparable measurements with X-rays and γ-rays from radon and ^{60}Co.

Table 4. *Inherent unsharpness, U_i, using Crystallex film and lead foil screens*

^{192}Ir: ¼ in. steel filter	0·17 mm
^{192}Ir: ½ in. steel filter	0·17 mm
360 kV: ⅛ in. lead filter	0·10 mm
1000 kV: ¼ in. steel filter	0·17 mm
Radon: ½ in. steel filter	0·29 mm
^{60}Co: ½ in. steel filter	0·36 mm

It will be noted that, as was found when considering the value of the absorption coefficient, there is again a good correspondence between ^{192}Ir radiation and 1000 kV X-rays. The inherent unsharpness with ^{192}Ir is considerably smaller than with other γ-ray sources such as radon and ^{60}Co, and ^{192}Ir should therefore give considerably better definition than these.

In considering the problem of unsharpness in a radiograph it is important to take into account the total unsharpness. This is usually made up of two components, the inherent and the geometric unsharpnesses.

Geometric unsharpness, arising from the finite size of the source of radiation, is given by the expression:

$$U_g = b \,.\, s/a \qquad (2)$$

where b = object-film distance; a = source-object distance; and s = source width, perpendicular to direction of rays.

$(b + a)$ represents the source-film-distance (S-F-D), which within the limits of exposure time has to be chosen by the radiographer. It is therefore necessary to decide how small the geometric unsharpness can be made. Unsharpnesses are not directly additive,[8,9,10] and a number of different formulae have been suggested for the summation of geometric and inherent unsharpnesses. That put forward by the author indicates that if

$U_g = 2U_i$, the total unsharpness $U_{total} = 2{\cdot}25\ U_i$
$U_g = U_i$, the total unsharpness $U_{total} = 1{\cdot}46\ U_i$
$U_g = \frac{1}{2}U_i$, the total unsharpness $U_{total} = 1{\cdot}13\ U_i$

The reduction in total unsharpness by making U_g less than half U_i is therefore very slight, and it is suggested that $U_g = \frac{1}{2}U_i$ be taken as the limiting value. Fig. 6 shows the values of source-film-distance, calculated for various source-diameters, for one type of fine-grain film. It is obvious that for the larger specimen thicknesses these distances would necessitate uneconomic exposure times, but they are included to emphasize that the larger the S-F-D's used, up to these values, the better will be the definition obtained.

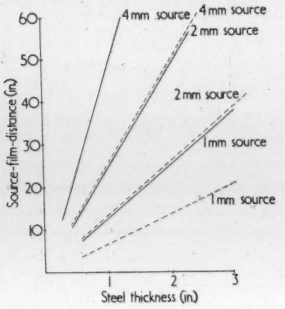


Fig. 6. Optimum source-film-distance against steel thickness, calculated for ————, $U_g = \frac{1}{2}U_i$; – – – –, $U_g = U_i$

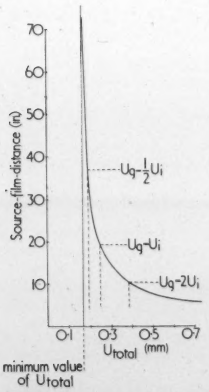


Fig. 7. Total unsharpness against source-film-distance, showing the effect of varying the relation between U_g and U_i, for one particular specimen thickness. ($1\frac{1}{2}$ in. steel)

If it is assumed that $U_g = U_i$ there is an increase in the total unsharpness of approximately 30%, but this condition does allow considerably shorter S-F-D's to be used: the broken lines on Fig. 6 show the optimum values of S-F-D calculated on this assumption. These shorter S-F-D's lead to a loss in definition and Fig. 7 shows a graph of S-F-D against total unsharpness, for one particular set of radiographic conditions, which illustrates the rapid rise in total unsharpness as the S-F-D is decreased. The points at which $U_g = \frac{1}{2}U_i$, $U_g = U_i$ and $U_g = 2U_i$ are indicated.

EXPERIMENTAL COMPARISON OF QUALITY WITH IRIDIUM 192 AND X-RAYS

In order to check the importance of radiographic unsharpness in its effect on the quality of a radiograph, a series of comparison radiographs were taken with ^{192}Ir and X-rays. The specimen was a steel strip containing a large number of fine cracks of different depths and widths, which was built up to different thicknesses. On this specimen the differences in radiographic quality were probably due to a combination of three effects:

(1) the loss in contrast due to the change in absorption coefficient;
(2) the loss in definition with ^{192}Ir radiation, due to the greater inherent unsharpness;
(3) the loss in contrast with ^{192}Ir on narrow defects such as cracks, due to the greater inherent unsharpness. This effect has been described in another paper;[8] it can be regarded as due to the interaction of unsharpness at the two edges of the image of the flaw, and its magnitude depends on the width of the flaw. With fine cracks, such as exist in this specimen, the effect was expected to be considerable.

The specimen was radiographed with X-rays and with ^{192}Ir radiation on the same type of film and the exposures adjusted so that the same density, and therefore the same film-contrast, was obtained. The X-radiographs were taken with an S-F-D of 40 in., and a kilovoltage chosen to require an exposure of about 20 mA-minutes. Description of the relative quality of the films is difficult, and reproductions from the radiographs, unless full-size, are liable to disguise much of the difference in quality, so that comparative prints are not included in this paper. On the $\frac{1}{2}$ in. specimen the obliteration of the fine shallow cracks, due to image spread produced by unsharpness, is quite easy to see, although the apparent blurring of the wider cracks appears to be much the same in the two cases. When the thickness is increased to 1 in. the same effect is noticeable; the finer cracks lose their sharp crack-like appearance on the ^{192}Ir radiograph, and the very fine cracks disappear altogether. Through $1\frac{1}{4}$ in. steel the contrast-differences are not so great, but still a number of the finer cracks, detectable by X-rays, cannot be shown with ^{192}Ir.

TECHNIQUE: FILM AND SCREEN CHARACTERISTICS

The importance of film gradient has already been emphasized and, in practice, this usually means the use of a fine-grain film of the type suitable for use with metal foil screens, exposed so as to have, after normal processing, a density of about 2·5–3·0. There is another film on the market, however,

not of the fine-grain class, which appears to have contrast characteristics, making it suitable for some γ-ray work. This is Ilford Industrial G, which has been found to have a contrast gradient almost as high as fine-grain type films and is of very high speed—six to ten times faster than fine-grain film. The film is very grainy, but the measured inherent unsharpness, which with most X-ray films has been found to vary directly with film speed, is not as high as would be expected for such a fast film, and the detail which the film is capable of showing is remarkably high. The use of this film for steel casting inspection with ^{192}Ir on the larger specimen thicknesses, where the length of exposure time makes it impractical to use a large S-F-D with fine-grain film, seems likely to give very satisfactory results. An exposure chart is shown in Fig. 8.

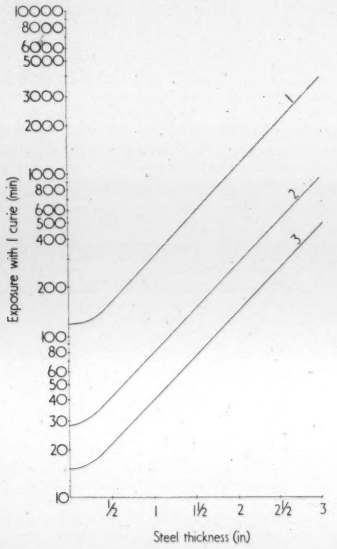


Fig. 8. Exposure chart for ^{192}Ir radiation, with steel

Curve 1, Crystallex, 20 in. source-film-distance; curve 2, Crystallex, 10 in. source-film-distance; curve 3, Industrial G, 20 in. source-film-distance. Lead screens, 0·004 in. front; 0·01 in. back; film density, 2·0; development, D-19b, 5 min, 68° F.

The optimum thickness of lead foil intensifying screens has been investigated. So far as the intensifying action is concerned the optimum thickness of the front screen is about 0·004 in., but is not at all critical. Any increase of thickness above 0·010 in. acts only as a filter and reduces the effectiveness of the screen in shortening the exposure-time. With the back screen there is no observable difference in intensification so long as the back screen is thicker than 0·002 in. Experiments have shown, however, that this screen assists in absorbing back scatter from material behind the film cassette as, for example, when it is necessary to have the cassette laid on a concrete floor and the optimum thickness for this purpose has been found to be 0·010–0·020 in. The recommended lead screen thicknesses for use with ^{192}Ir are, therefore, 0·004 in. front and 0·010 in. back.

THICKNESS LATITUDE

It is frequently necessary to radiograph a specimen in which the thickness varies considerably, and it remains to discuss how the sensitivity varies over different ranges of thickness. It will be assumed that the thinnest part of the specimen is recorded by a film density of 3·5 (the limit set by most good film illuminators). Fig. 9 shows the calculated

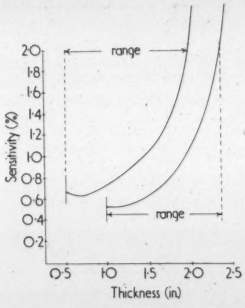


Fig. 9. Graph showing the relation between thickness latitude and sensitivity with ^{192}Ir radiation

change in sensitivity with thickness for two specimens, having as their thinnest sections ½ and 1 in. respectively. If a sensitivity of 2% be taken as the lowest acceptable limit, this gives the thickness ranges which can be examined on one radiograph as 0·5–1·9 in. and 1·0–2·4 in. in the two examples chosen. In practice, this range can often be extended further by the use of a "sandwich" of two films of different speeds and three lead foil screens; the thicker portions of the specimen are then recorded on the fast film and the thinner parts on the slower one, without the necessity for two separate exposures.

CONCLUSIONS AND RECOMMENDATIONS

The experimental and theoretical estimates of sensitivity have shown that, compared with 200 kV X-rays the sensitivity obtainable with ^{192}Ir is markedly inferior and it is suggested that for the detection of defects such as fine cracks, ^{192}Ir should not be used when the use of X-rays is practicable. Even with steel thicknesses of 2 in. and more, when 400 kV X-rays would normally be used, X-ray techniques are superior for the detection of this type of defect, though in this region the difference in penetrameter-sensitivity is small.

An estimate of the equivalent X-ray voltage has been made and suggests that equal sensitivity ought to be obtained with 800–900 kV X-rays.

It is then to that class of inspection where critical definition is not so important that ^{192}Ir is likely to be most suitably applied. It has been shown that the flaw-sensitivity, when considered in terms of contrast alone, is comparable with 400 kV X-rays for thicknesses from 1½–3 in. steel, and that the sensitivity obtainable with ^{192}Ir radiation, although considerably less than that with 200 kV X-rays, remains reasonable for thicknesses down to about ½ in. steel, provided

that film is used under such conditions that a film-gradient of four or greater is obtained.

With the above reservation, ^{192}Ir radiation is considered suitable for the inspection of steel castings over the range of thickness from $\frac{1}{2}$ to 3 in.

For the inspection of welds, where small fine cracks and narrow slag lines are likely to be prevalent, such as welding in high-alloy steels, ^{192}Ir is not suitable and should not be used to replace X-radiography. For a considerable class of welding, however, where very fine defects other than cracks would not normally be considered to be serious, the sensitivity obtainable with ^{192}Ir radiation is thought to be adequate, particularly on the thicker sections, and in view of the rapid fall-off in sensitivity on thin sections, already discussed, it is suggested that a steel thickness of $\frac{3}{4}$ in. be taken as a minimum for this class of work.

ACKNOWLEDGEMENTS

Acknowledgement is made to the Chief Scientist, Ministry of Supply, for permission to publish this paper, and to colleagues at Woolwich and Ruddington for assistance in its preparation. Crown copyright reserved. Reproduced with the permission of the Controller of H.M. Stationery Office.

REFERENCES

(1) A.E.R.E., Harwell. *Data sheets on artificial radio-isotopes* (London: H.M. Stationery Office, 1950).
(2) MORRISON, A. *Nat. Res. Coun. Canada Rep. No.* 2407. (1951).
(3) MÜLLER, D. E., HOYT, H. C., and others. *Phys. Rev.*, **88**, pp. 775–93. (1952).
(4) ROULSTON, K., PRINGLE, R. *Phys. Rev.*, **87**, pp. 930–1 (1952).
(5) VAN DE GRAAFF, R. J. *A.S.T.M. Bulletin*, pp. 54–64 (Dec., 1948).
(6) *X-Ray exposure data-sheets* (London: Newton-Victor Ltd., 1948).
(7) KRUITHOF, A. M. *Philips Tech. Rev.*, **11**, p. 333 (1950).
(8) HALMSHAW, R. Unpublished report (1951).
(9) BERTHOLD, R. *Atlas of Non-destructive Testing* (Leipzig: J. A. Barth, 1938).
(10) KLASENS, H. A. *Philips Res. Rep.*, **1**, p. 241 (1946).

The critical evaluation of compression data for liquids and a revision of the isotherms of mercury

By K. E. BETT, B.Sc., K. E. WEALE, Ph.D., and D. M. NEWITT, D.Sc., F.R.S., Imperial College of Science and Technology, London, S.W.7

[*Paper received* 23 *September*, 1953]

It is shown that Hudleston's equation represents liquid compression data at least as accurately as Tait's equation, and the straight line isotherms obtained are more convenient for the estimation of random experimental errors. The equation may also be used to extrapolate and interpolate measurements over a large P–T region.

Revised values for the isothermal compression of mercury from -30 to 150° C and up to 12 000 kg/cm^2 have been obtained by using the equation to correlate existing measurements with more recent sound-velocity determinations. Some thermodynamic derivatives for mercury are also calculated.

Equations of state for liquids may take the form of convergent series equations which require several virial coefficients related in an unknown way to molecular force fields. An alternative approach by way of statistical mechanics depends upon simplifying assumptions as to the structure of groups of molecules in the liquid; and equations derived in this way are at present restricted by mathematical and computational difficulties to the idealized case of spherically symmetrical force fields. Neither approach is, therefore, of immediate practical value and empirical equations are accordingly used for correlating experimental data over restricted ranges of temperature and pressure, and for interpolation and extrapolation.

The purpose of the present paper is to show how one such equation may be used to correlate P–V–T data for mercury and some typical liquids over a wide range of pressure and temperature and to assess their consistency. The equation can also be used to relate the values of the derivative $(\partial V/\partial P)_T$ at high pressures to that at atmospheric pressure and to calculate the principal thermodynamic functions of a liquid.

REPRESENTATION OF COMPRESSION DATA AS A FUNCTION OF PRESSURE

If the specific volume of a liquid at any temperature is plotted against the pressure, a smooth curve is obtained which, for most liquids, can only be represented over an extended pressure range by a polynomial equation of at least the fourth degree. For limited pressure ranges up to, say, 2000 atm, quadratic equations containing arbitrary temperature-dependent constants will correlate the compression data satisfactorily, but such equations are not of general validity.

As van der Waals' equation undoubtedly represents in a general way the behaviour of a liquid over a limited pressure range, many equations, such as those proposed by Tumlirz,[1] Macleod[2] and Brinkman,[3] have been based upon it.

One of the earliest empirical equations is that due to Tait[4]:

$$k = C \log\left(\frac{B+P}{B}\right) \qquad (1)$$

where k is the compression at pressure P, and B and C are constants, the former usually being a quadratic function of temperature and the latter independent of temperature.

Gibson and his associates[5,6] have made a thorough study of this equation and find that it fits closely their data for a number of liquids up to 1000 bars over a temperature range of 25 to 85° C, and Adam's [7] data for water at 25° C up to 10 000 bars.

It will be noted that Tait's equation leads to zero specific

volume at a finite high pressure; recently Kirkwood has modified it to remove this anomaly, as follows:

$$\log\left[\frac{v_p}{v_o}\right]_S = \frac{1}{n}\log\left(\frac{A_S + P}{A_S}\right) \qquad (2)$$

where A_S is a function of the entropy, $1/n$ is a constant corresponding to C in the original equation, and v_p and v_o are volumes measured along an adiabatic, the subscripts relating to pressure.

This equation has been fitted to Bridgman's[8] data for water by Richardson, Arons and Halverson,[9] who found deviations of less than 4% in n over a large P–V field extending to 25 000 kg/cm^2.

Various equations of a similar type are to be found in the literature and attempts have been made to identify their constants with properties of simplified intermolecular force fields and to relate them to viscosity and other data (e.g. Hudleston,[10] Macleod[2]).

Of these, we have found Hudleston's equation to be the most useful for correlation purposes. It is derived from a postulated law of intermolecular force

$$F = \lambda(L_0 - L)\exp\left[\theta(L_0 - L)\right] \qquad (3)$$

in which F is the force acting on one face of a cube containing unit mass of the liquid, L is the length of the cube side, and λ and θ are specific constants. L_0 is the value of L at zero (or unit) pressure.

Noting that $F = L^2P$, then

$$\log_{10}\frac{L^2P}{L_0 - L} = A^T + B(L_0 - L) \qquad (4)$$

where $A^T = \log_{10}\lambda$ and $B = \theta\log_{10}e$. It follows from this equation that $\log_{10}\dfrac{L^2P}{L_0 - L}$ should be a linear function of $(L_0 - L)$.

In order to apply the relationship to the P–V isotherms of a particular liquid, the values of L and L_0 have to be evaluated from the compression and thermal expansion data.

We can write $L = (V_P^T/V_0^t)^{\frac{1}{3}}$ and $L_0 = (V_0^T/V_0^t)^{\frac{1}{3}}$, where V is the volume of a given mass of liquid, the superscript denoting the temperature and the subscript the pressure. Subscript zero signifies atmospheric pressure and the superscripts T and t denote the temperature of an isotherm and some reference temperature, usually 0° C.

The isothermal compression, $k = (V_0^T - V_P^T)/V_0^T$

Thus $$L = L_0(1 - k)^{\frac{1}{3}}$$

Hudleston[10] applied the equation to the data of Bridgman[11] over the pressure range 1000 to 12 000 atm, and found that a linear relationship existed for "normal" liquids, such as carbon disulphide, ether and phosphorus trichloride, but not for "abnormal" liquids, such as water and acetone. Since he did not state any criterion for "normality" and "abnormality," little use has hitherto been made of his equation.

In the course of our examination of compression data, we have found that the equation is of general application and that it holds for both water and acetone to within the experimental error of the original measurements.

APPLICATION OF HUDLESTON'S EQUATION TO THE CORRELATION OF LIQUID COMPRESSION DATA

The compressions used by Hudleston to test his equation are based on Bridgman's[12] measurements using a piston displacement type of piezometer. The data for some of the liquids only obey the relationship to within the estimated experimental error at pressures above about 2500 kg/cm^2; below this pressure, the deviation from a straight line increases. For example, the compression of acetone at 20° C and 1000 kg/cm^2 calculated from Hudleston's equation differs from the observed value by as much as 4·8%. Such deviations at low pressures led Hudleston to believe that the relationship failed for certain liquids. The piston displacement piezometer is known to be inaccurate at low pressures and Bridgman found it necessary to supplement his measurements at 500 kg/cm^2 with values obtained from Amagat's work. If Amagat's[13] data for acetone to 3000 atm and Newitt and Weale's[14] data up to 1000 atm are plotted, using Hudleston's co-ordinates, straight lines are obtained, showing that the compressions of acetone do conform to the relationship below 2500 kg/cm^2.

Although Hudleston's relationship will confirm the consistency of a series of compression measurements and reveal the extent of random experimental errors, the fact that the results obey it is not a criterion of their absolute accuracy. For instance, both Amagat's[13] and Bridgman's[12] values for the compression of ethyl ether at 20° C obey the relationship well, except for the value which Bridgman assumed at 500 kg/cm^2; nevertheless, the values at 1000 kg/cm^2 differ by 5·7%.

In order to investigate the effect of change in temperature on the Hudleston constants, A and B, the relationship has been applied to the most reliable compression data available. Gibson and his co-workers[6,15,16,17] have made very accurate measurements on pure liquids at pressures up to 1000 bars from 25 to 85° C, using a re-entrant capillary piezometer made of quartz or vitreous silica. Their compression results are presented in the form of the Tait equation, but, by solving this equation and correcting the calculated compressions in accordance with the tables of deviations published in their papers, the experimentally observed values can be calculated and used to derive the Hudleston parameters.

This has been done for water, carbon tetrachloride and benzene. Since the expansion of water at atmospheric pressure was not measured by Gibson,[17] the data needed to calculate $L_0 = (V_0^T/V_0^0)^{\frac{1}{3}}$ have been taken from the revised Chappius table published by Tilton and Taylor.[18] In the case of benzene[6] and carbon tetrachloride[17], the observed specific volumes are in good agreement with the recent values reported by Timmermans[19] and have been used to calculate $L_0 = (V_0^T/V_0^{25})^{\frac{1}{3}}$. The reference temperature is taken at 25° C, as the freezing point of each liquid is above 0° C.

In Figs. 1 and 2, the mean experimental values for the compression of water and carbon tetrachloride at any one pressure and temperature have been plotted using the Hudleston co-ordinates. The random experimental error in Gibson's results is so small that it proved to be impossible to separate all the individual compression measurements on the graphs; the point for water at 1000 bars and 25° C, for instance, represents the mean of 21 independent measurements.

It can be seen from these diagrams that the straight line isotherms are approximately parallel. Additional lines show the relationship between the temperature (right hand ordinate) and the intercepts of Hudleston's isotherms on the logarithmic axis.

The intercepts and the slopes of the isothermal Hudleston relationships which have been fitted to all the experimentally determined compressions for each liquid by the method of least squares are given in Table 1.

Table 1. *The intercepts A and slopes B of the Hudleston isotherms for water, carbon tetrachloride, and benzene*

T °C	Water		Benzene		Carbon tetrachloride	
	A	B	A	B	A	B
25	4·8194	3·41	4·4905	5·935	4·4464	6·103
35	4·8265	3·55	4·4592	5·933	4·4118	6·163
45	4·8316	3·43	4·4222	5·982	4·3799	6·067
55	4·8334	3·33	4·3890	5·997	4·3441	6·102
65	4·8288	3·44	4·3542	5·965	4·3088	6·061
75	4·8221	3·51	—	—	—	—
85	4·8134	3·43	—	—	—	—
	Mean	3·44		5·962		6·099

CHARACTERISTICS OF THE PARAMETERS OF HUDLESTON'S EQUATION

(*a*) *Slopes of the isotherms.* Gibson[6,17] states that the experimental error in his compression data for benzene and carbon tetrachloride reduce to the uncertainty in the pressure measurement, which is ±0·5 bar. This corresponds to a possible maximum error of ±0·07 in the slope of a Hudleston isotherm. The observed maximum deviation between the slope of any isotherm and the mean is ±0·06 and the average deviation is ±0·025. The average deviation of the slopes of the isotherms for water is similarly within the possible experimental error, which is slightly higher than for the other two liquids. Thus, for any one of the three liquids, the isotherms have identical slopes within the experimental accuracy.

(*b*) *Intercepts of the isotherms.* It is found that for most liquids the intercepts of the Hudleston isotherms on the logarithmic axis are a linear function of temperature, as shown for carbon tetrachloride in Fig. 1. The values for

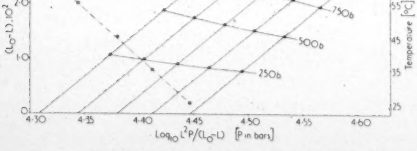


Fig. 1. Hudleston diagram for carbon tetrachloride

benzene and carbon tetrachloride given in Table 1 can be represented by the following equations over the temperature range 25 to 65° C:

For carbon tetrachloride:

$$A^T = 4{\cdot}5325 - 0{\cdot}00343\,T \qquad (5)$$

For benzene:

$$A^T = 4{\cdot}5772 - 0{\cdot}00343\,T \qquad (6)$$

The average deviation of A from that calculated by these equations is only ±0·17% for both liquids.

A property of Hudleston's equation is that at low pressures, when $(L_0 - L) \simeq 0$, then $L^2P/(L_0 - L) \simeq 3\,L_0/\beta_0^T$.

Thus, the intercept A is given by

$$A^T = \log_{10}(3L_0/\beta_0^T) \qquad (7)$$

where β_0^T is the coefficient of isothermal compressibility at atmospheric pressure and temperature, T° C.

Unfortunately, precise measurements of the isothermal compressibility coefficients are difficult to make at low pressures and thus the intercept A, at various temperatures, cannot be accurately calculated. In order to overcome the experimental difficulties inherent in the direct measurement of this coefficient, volumetric methods for determining the adiabatic compressibility have been devised by Tyrer[20,21], Villey[22] and Philip.[23]

The isothermal coefficient can then be calculated from the adiabatic compressibility by means of the relation:

$$\beta_0^T = \beta_0^S + \frac{\tau p}{C_p}\left(\frac{\partial v}{\partial T}\right)^2 \qquad (8)$$

where β_0^S is the adiabatic compressibility coefficient
C_p is the specific heat at constant pressure,
τ is the absolute temperature,
$\frac{\partial v}{\partial T}$ is the rate of change of specific volume with temperature,
and ρ is the density of the liquid.

More recently, accurate methods for measuring the velocity of sound in liquids at ultrasonic frequencies have been developed, which allow the adiabatic coefficient to be calculated to within about 0·1%. Thus, the isothermal coefficient β^T and the intercept A can be calculated, provided that the sound frequency is sufficiently below the dispersive region and the required thermodynamic data are known with sufficient precision. The necessary data for benzene are given in Table 2.

The velocity of sound measurements is the mean of the results of Freyer, Hubbard and Andrews,[24] and of Lagemann, McMillan and Woolf[25] (with a value at 60° C from Lagemann and others). The maximum deviation of either set of results from the mean at any temperature is ±0·12%.

The density from 20 to 60° C is taken from the work of Cohen and Buij[26] and that at 10° C is the mean of all the recent determinations reported by Timmermans.[19]

The values of $\partial v/\partial T$ are calculated from the specific volume measurements of Patterson,[27] Young,[28] and Cohen and Buij,[26] and the specific heat at constant pressure is the mean of values reported by Tréhin,[29] Huffman, Parks and Daniels,[30] Stull[31] and Oliver, Eaton and Huffman.[32]

It is found that the calculated intercepts agree with those given by equation (6) derived from P–V measurements to within ±0·07%. Similar agreement is found in the case of carbon tetrachloride and water, showing that Hudleston's relationship will satisfactorily correlate the isothermal compressibility coefficient at atmospheric pressure obtained from ultrasonic data, with the compressions obtained at high pressure.

Table 2. *Calculation of the initial coefficient of compressibility of benzene from velocity of sound data*

T °C	u m.s⁻¹	ρ g.cm⁻³	∂v/∂T cm³g⁻¹ °C⁻¹	C_p joule.g⁻¹ °C⁻¹	$\beta_0^T 10^6$ bars⁻¹	A
10	1374·5	0·8894	0·001 30	0·403	84·9	4·5457
20	1325·5	0·8791	0·001 36	0·410	92·4	4·5105
30	1279·0	0·8684	0·001 42	0·418	100·5	4·4758
40	1232·5	0·8577	0·001 47	0·426	109·4	4·4410
50	1184·0	0·8469	0·001 53	0·434	119·4	4·4048
60	1143·0	0·8360	0·001 58	0·441	129·3	4·3719

It can be seen from Fig. 2 that water behaves abnormally in that A passes through a maximum at about 50° C, the temperature of maximum initial compressibility. For most liquids, however, the isotherms are equidistant and parallel,

so that the P–V–T data may be accurately correlated by an equation of the type

$$\log_{10}\frac{L^2P}{L_0-L} = A^\circ - CT + B(L_0 - L) \qquad (9)$$

For benzene from 25 to 65° C, at pressures up to 1000 bars,

$$A^\circ = 4{\cdot}5772, \quad B = 5{\cdot}962, \quad C = 0{\cdot}00343$$

and, for carbon tetrachloride over the same P–T region,

$$A^\circ = 4{\cdot}5325, \quad B = 6{\cdot}099, \quad C = 0{\cdot}00343$$

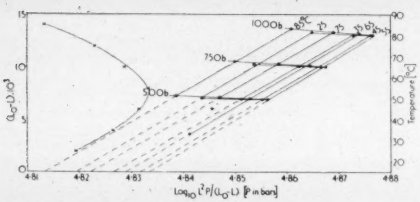


Fig. 2. Hudleston diagram for water

COMPARATIVE ACCURACY OF TAIT'S AND HUDLESTON'S EQUATIONS

If the Hudleston isotherms are solved and the compressions evaluated at 250, 500, 750 and 1000 bars for each temperature, the calculated compressions may be compared with the observed ones and a deviation table constructed. Gibson's charts showing the deviations of the compressions given by the Tait equation from the experimental values allow a comparison to be made between the deviations from both empirical equations. This has been done for carbon tetrachloride in Table 3, in which D denotes the deviation (observed minus calculated compression) $\times 10^5$ for both Tait's and Hudleston's equations.

Table 3. *Comparison of the accuracy of correlation of the carbon tetrachloride compressions given by Tait's and Hudleston's equations*

T	ΣD		ΣD^2	
°C	*Tait*	*Hudleston*	*Tait*	*Hudleston*
25	+5	+2	197	90
35	+6	+1	70	11
45	+7	−4	111	121
55	+5	+3	203	181
65	0	−1	104	19

Thus Hudleston's equation correlates the observed compressions for carbon tetrachloride somewhat better than Tait's equation. An examination of Gibson's data for benzene and for water has given similar results.

The isothermal compressibility coefficients for benzene at atmospheric pressure have been calculated from both Tait's and Hudleston's equations, and are given in Table 4, together with those values obtained from velocity of sound data.

The coefficients obtained from the extrapolated Hudleston isotherms are in better agreement with those obtained from ultrasonic data than are those calculated from Tait's equation. It is, however, at high pressures that the two equations differ appreciably. The compressions calculated from both Tait's and Hudleston's isotherms, which have been fitted to Gibson's data for benzene at 25° C, agree to within 0·12% over the experimental pressure range 250 to 1000 bars. Extrapolating beyond this range to 5000 bars, the discrepancy is 1·7% and, at 10000 bars, it is 2·6%. Ultimately the equations diverge widely because Tait's equation requires zero specific volume at a pressure of approximately 4×10^7 bars, while according to Hudleston's equation this limit is approached asymptotically at infinite pressure.

Table 4. *Comparison of the initial isothermal compressibility coefficients for benzene with those obtained from Tait's and Hudleston's equations*

T	*Initial coefficient* $10^6\beta^T$ *bars*$^{-1}$		
°C	*Sound velocity*	*Hudleston*	*Tait*
10	84·9	85·4	86·5
20	92·4	92·8	93·1
30	100·5	100·8	100·4
40	109·4	109·6	108·6
50	119·4	119·1	117·8
60	129·3	129·4	128·2

COMPRESSION OF ETHYL ALCOHOL

To show the use of Hudleston's equation in correlating compression data from different sources, we have applied it to ethyl alcohol. The properties of Hudleston's isotherms, derived from the consideration of Gibson's results, over a limited P–T region, are found to hold for ethyl alcohol over

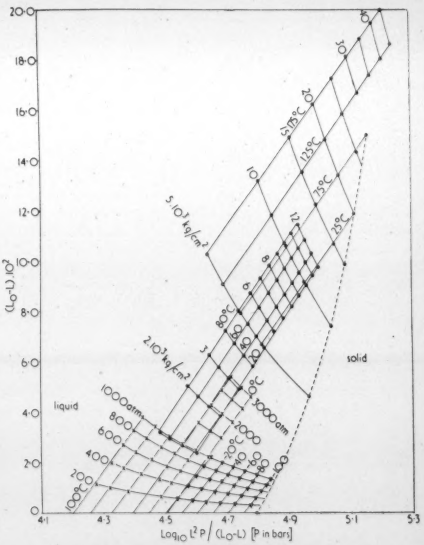


Fig. 3. Hudleston diagram for ethyl alcohol
× Amagat (1893); ▽ Seitz and Lechner (1916); ○ Bridgman (1913, 1942)

the much more extensive region for which data have been obtained for this liquid.

In Fig. 3 the data are plotted using Hudleston's coordinates, the base volume for L_0 being taken at 0° C and atmospheric pressure. The different pressure units used in the original investigations are shown in this figure but these units have been converted to bars for the evaluation of the logarithmic term. The compressions below 1000 atm, from 0 to −100° C, were measured by Seitz and Lechner[33] and those from 0 to 100° C by Amagat.[13] Both sets of data fall on approximately equi-spaced and parallel isotherms, but the slope of Amagat's isotherms differs from that of Seitz and Lechner's; the maximum deviation in the compression at 0° C and 1000 atm is 2%. Amagat also measured the compression of ethyl alcohol by a different method up to 3000 atm in the temperature interval 0 to 40° C. These results agree with his other values to within 1% at 1000 atm.

Above 3000 atm the only compressions determined are those due to Bridgman.[12,34] The earlier results cover a temperature range from 20 to 80° C, and pressures up to 12 000 kg/cm². Because of uncertainties in Bridgman's data at low pressures, the isotherms have not been drawn below the 1000 kg/cm² isobar. At temperatures below 80° C, the divergence between Amagat's and Bridgman's results is less than 0·5% at 1000 atm, but increases to about 2% at 3000 atm and 20° C. The second set of results covers a pressure range from 5000 kg/cm² to 45 000 kg/cm² at temperature intervals of 50° C from 25 to 175° C. These values are based on the compression at 5000 kg/cm² obtained from the earlier work and are in no sense absolute, but they are satisfactorily represented by Hudleston's equation up to 45 000 kg/cm² pressure.

The diagram shows clearly how the equation can be applied to correlating and interpolating compression data over a large *P–T* region. In Table 5 the constants *A* and *B* of the isotherms, fitted to the experimental data of each investigator by the method of least squares, have been tabulated. Values of *A* calculated from the adiabatic coefficient of compressibility measured by Tyrer,[21] using a volumetric method, and by Freyer, Hubbard and Andrews,[24] using a "sonic interferometer," are given in Table 6. The specific volume data needed to calculate *L*, and hence *A*, are the mean of the recent determinations reported by Timmermans.[19]

The intercepts *A*, in Table 6, calculated from the adiabatic compressibility coefficients, and those in Table 5, obtained from the extrapolated Hudleston isotherms, can be represented by the following equation to within 0·6% over the temperature range −100 to +100° C.

$$A^T = 4{\cdot}500 - 0{\cdot}00267\,T$$

The isotherms for each set of data all have approximately the same slope, but the average slope of each group is different; the discrepancy being a measure of the systematic error inherent in each method. In the absence of such errors, the larger the pressure range covered the more accurately the slope is defined. If the true slope of the isotherms is then taken as the mean of Bridgman's results and those of Amagat up to 3000 atm, the *P–V–T* data for the liquid can be represented by the equation

$$\log_{10}\frac{L^2P}{L_0 - L} = 4{\cdot}500 - 0{\cdot}00267T + 5{\cdot}79(L_0 - L)$$

This equation gives the experimentally observed compressions of ethyl alcohol at all pressures and temperatures from −100 to +175° C and up to 45 000 kg/cm², with an accuracy of ±5%, which is within the experimental error of Bridgman's results at the higher pressures. It can also be used to predict the compression of ethyl alcohol in areas where no results are available, such as the region above 1000 atm pressure and below 0° C. The extrapolation of results by Hudleston's relationship is not possible if a phase change occurs in the region under consideration, since the accompanying volume change causes a "step" in the isotherms. In Fig. 3 the freezing parameters measured by Bridgman[34] have been used to construct the solid-liquid equilibrium curve.

Table 5. *The intercept A and the slope B of Hudleston's isotherms for ethyl alcohol*

T °C	*Amagat I* 0–1000 atm *A*	*B*	*Amagat II* 0–3000 atm *A*	*B*	*Bridgman I* 1000–12 000 kg.cm⁻² *A*	*B*
0	4·494	6·18	4·504	5·90	—	—
20	4·436	6·34	4·444	6·08	4·470	5·40
40	4·383	6·23	4·394	5·91	4·407	5·56
60	4·324	6·16	—	—	4·349	5·60
80	4·256	6·08	—	—	4·298	5·53
100	4·200	6·01	—	—	—	—
mean	—	6·17	—	5·96	—	5·52

T °C	*Seitz and Lechner* 0–1000 atm *A*	*B*	*T* °C	*Bridgman II* 5000–45 000 kg.cm⁻² *A*	*B*
0	4·501	6·59	25	4·422	5·86
−20	4·551	6·58	75	4·277	5·90
−40	4·600	6·53	125	4·169	5·76
−60	4·652	6·20	175	4·041	5·83
−80	4·701	6·28	mean	—	5·84
−100	4·751	6·49			
mean	—	6·45			

Table 6. *Calculation of the intercept A from static and dynamic determinations of the adiabatic compressibility coefficient of ethyl alcohol*

T °C	*Adiabatic compressibility coefficient* $10^6\beta S$ bars⁻¹ Tyrer	F.H.A.	*Isothermal compressibility coefficient* $10^6\beta T$ bars⁻¹ Tyrer	F.H.A.	*Hudleston intercept* *A* Tyrer	F.H.A.
0	82·8	80·4	98·6	96·3	4·483	4·494
10	88·3	86·5	104·9	103·0	4·458	4·466
20	94·3	92·9	111·6	109·8	4·433	4·440
30	100·9	99·6	119·0	118·0	4·406	4·410
40	108·2	106·8	127·4	126·1	4·378	4·383
50	116·3	115·1	137·0	136·0	4·348	4·352
60	125·0	—	147·4	—	4·318	—
70	135·2	—	159·7	—	4·286	—

COMPRESSION OF MERCURY

Mercury is extensively used in high pressure work, both as a pressure transmitting fluid and as the pressure element in certain types of standard primary gauge. It is, therefore, of importance to know its compression with the greatest possible accuracy.

At present, the only results covering a large pressure range are those obtained by Bridgman[35] at 0 and 22° C and pressures up to 12 000 kg/cm². We have re-examined these data, using Hudleston's relationship to assess their consistency and to correlate them with measurements of the isothermal compressibility coefficients at atmospheric pressure.

Bridgman's results at 0° C were obtained by weighting and smoothing ninety-one separate determinations made with five

different piezometers, and those at 22° C by correlating thirty-eight results obtained with the two piezometers which gave the most consistent measurements at 0° C. The resulting compressions have been plotted in Fig. 4, in which the base

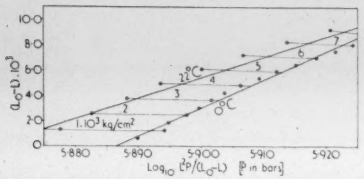


Fig. 4. Bridgman's values for the compression of mercury at 0 and 22° C, as represented by the Hudleston relationship

volume for L_0 and L is taken at 0° C and the values of L_0 calculated from Sear's thermal expansion formula.[36] As with ethyl alcohol, Bridgman's pressure unit has been converted to bars for the evaluation of the logarithmic term. The straight line isotherms indicate that some of the results at low pressure are in error by as much as 0·5%. That these deviations are the result of the smoothing procedure can be shown by fitting the Hudleston relationship to the unsmoothed experimental results obtained with each piezometer.

Fig. 5, showing the results obtained at 22° C with piezometer No. 10, is typical of all the piezometers. In general,

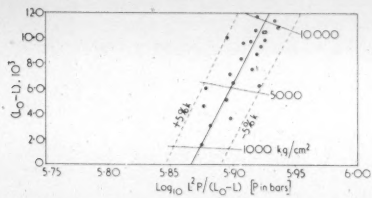


Fig. 5. Hudleston relationship for Bridgman's compression results at 22° C obtained with piezometer No. 10

the average random variation of the data about the isotherms is ±2% of the compression, but it tends to increase at lower pressures; there is, however, no evidence to suggest that the isotherms are non-linear. A few results, such as the one in Fig. 5 and five obtained with piezometer No. 8 at 22° C, are widely divergent; some of these have been attributed to a permanent distortion of the piezometer, due to the water freezing at high pressure. These points have been discarded by fitting a straight line to all the results for each piezometer and then rejecting those points which lie outside the arbitrarily chosen fiducial limit of ±5% of the compression; a final isotherm has then been fitted to the remaining compressions. By this procedure, a total of nine points were rejected at 0° C and six at 22° C.

The intercept A, the slope B of the various isotherms, the number of compression results correlated N, and the mean values of the two Hudleston co-ordinates $\bar{X}=1/N\times\Sigma_1^N[\log_{10} L^2P/(L_0-L)]$ and $\bar{Y}=1/N\times\Sigma_1^N(L_0-L)$, are given in Table 7. The best correlation of Bridgman's results at either temperature is given by the isotherms defined by the mean slope and the mean intercept and passing through the mean value of $\bar{X}$ and $\bar{Y}$. Contrary to the findings for other liquids, the mean slope of the isotherm at 0° C is not

Table 7. *The Hudleston parameters for Bridgman's mercury compressions obtained with each piezometer*

T °C	Piezometer No.	N	$\bar{X}$	$\bar{Y}$	Intercept A	Slope B
0	4	16	5·9092	0·004 652	5·883	5·55
0	10	8	5·8994	0·004 837	5·886	2·76
0	8	22	5·9189	0·004 984	5·892	5·47
0	8A	12	5·9117	0·004 934	5·894	3·51
0	11	24	5·8954	0·004 645	5·868	5·97
		mean	5·9069	0·004 810	5·885	4·65
22	10	23	5·9082	0·007 947	5·867	5·22
22	8	9	5·9080	0·006 775	5·866	6·17
		mean	5·9081	0·007 361	5·866	5·69

the same as that at 22° C. This, however, is not surprising, because the variation in slope for the different piezometers is so large. If it is assumed that the two slopes should be the same, and equal to the mean of all seven slopes, i.e. 4·95, then the isotherms can be adjusted so that they pass through the mean values of $\bar{X}$ and $\bar{Y}$, and hence the corrected intercepts can be calculated. These are, at 0° C, $A = 5{\cdot}8831$ and, at 22° C, $A = 5{\cdot}8717$. The coefficients of isothermal compressibility calculated from these intercepts are $3{\cdot}926\times10^{-6}$ bars^{-1}, and $4{\cdot}036\times10^{-6}$ bars^{-1}, compared with Bridgman's values, which, when converted to bars, are equivalent to $3{\cdot}875\times10^{-6}$ bars^{-1} and $4{\cdot}028\times10^{-6}$ bars^{-1} respectively. Although the agreement is within 0·2% at 22° C, the coefficient at 0° C is 1·3% smaller than that published by Bridgman. It is shown below that the coefficients calculated from the velocity of sound measurements in mercury are in closer agreement with those we have calculated than with Bridgman's values.

Hubbard and Loomis[37] measured the velocity of sound of frequency $0{\cdot}44\times10^6$ sec^{-1} in mercury at 10° C temperature intervals from 0 to 70° C; the average deviation of the results at 20° C being 0·02%. The only other values were obtained by Kleppa[38] at 50 and 150° C, but, owing to pulse distortion and loss of signal strength, their accuracy was only ±1%. Nevertheless, the velocity of 1440 ± 10 m sec^{-1} at 50° C agrees with that measured by Hubbard and Loomis to within the experimental error of the results. Since Ringo, Fitzgerald and Hurdle[39] have made velocity measurements over the frequency range 10^8 to 10^9 sec^{-1} without finding any evidence of dispersion, a comparison of the compressibility coefficients calculated from Hubbard and Loomis' data with those obtained from extrapolated P–V isotherms is justified.

The data needed to calculate the adiabatic and isothermal compressibility coefficients and the Hudleston intercepts for mercury are presented in Table 8.

The velocities of sound from 0 to 70° C are taken from the work of Hubbard and Loomis[37] and the value at 150° C is due to Kleppa.[38] The density of mercury as a function of temperature and the rates of change of specific volume with temperature are calculated from Sear's expansion formula,[36] assuming the density at 0° C is 13·595 10 g/cm^3. This formula has been confirmed by Harlow[40] and Beattie and his co-workers.[41] The specific heats at constant pressure are calculated from the results of Douglas, Ball and

Table 8. *Calculation of the initial coefficients of compression of mercury from velocity of sound data*

T °C	u m.sec^{-1} ±0·3	ρ g.cm^{-3}	$\partial\nu/\partial T$ cm^{3}g^{-1} °C^{-1} ×10^{-6}	Cp joule. g^{-1} °C^{-1} ±0·0001	β^S bars^{-1} ×10^{-7} ±0·015	β^T bars^{-1} ×10^{-7} ±0·015	A ±0·0002
0	1460·2	13·595	13·347	0·1405	34·50	39·21	5·8837
10	1455·6	13·570	13·361	0·1401	34·78	39·68	5·8789
20	1451·0	13·546	13·375	0·1397	35·06	40·15	5·8740
30	1446·4	13·521	13·390	0·1393	35·35	40·63	5·8691
40	1441·7	13·497	13·405	0·1390	35·65	41·11	5·8642
50	1437·1	13·473	13·421	0·1386	35·94	41·60	5·8593
60	1432·4	13·448	13·438	0·1382	36·24	42·09	5·8545
70	1427·7	13·424	13·456	0·1379	36·55	42·60	5·8496
150	1370 ± 10	13·231	13·650	0·1360	40·3 ± 0·6	48·0 ± 0·6	5·800 ± 0·005

Ginnings,[42] and Bushey and Giauque,[43] assuming that the atomic weight of mercury is 200·61 and defining one 15° C calorie as $4·185 \times 10^7$ erg.

The Hudleston intercept A, calculated from the velocity of sound data, is a linear function of temperature as expected and is given by

$$A^T = 5·8837 - 0·0004877T \qquad (10)$$

From which the coefficient of compressibility is calculated to be $3·921 \times 10^{-6}$ bars^{-1} at 0° C, and $4·024 \times 10^{-6}$ bars^{-1} at 22° C, compared with values of $3·926 \times 10^{-6}$ and $4·036 \times 10^{-6}$ obtained from our correlation of Bridgman's results. Thus, the agreement is within 0·3% at both temperatures.

If the slope of Hudleston's isotherms, as deduced from Bridgman's work, is combined with the relationship between the intercept and temperature obtained from the velocity of sound measurements, the following equation can be formulated:

$$\log_{10}\frac{L^2P}{L_0 - L} = 5·8837 - 0·0004877T + 4·95(L_0 - L)$$

This equation has been solved and the relative volumes of mercury at temperatures from −30 to 150° C and pressures up to 12000 kg/cm^2 are given in Table 9. The pressures have been tabulated in kg/cm^2 so as to facilitate comparison between this table and the one published by Bridgman[35] and the freezing pressures have been taken from the work of Michels, Wassenaar and Blaisse,[44] and Bridgman.[35]

The random experimental errors in the relative volumes have been estimated for the isotherms at 0, 50, 100 and 150° C, by combining the experimental error in the intercept A, calculated from the velocity of sound data, with the uncertainty in the slope B of the isotherms deduced from equilibrium P–V–T data. Over the temperature range 0 to 70° C, equation (10) represents the intercept as a function of temperature to within the estimated maximum experimental error of ±0·0002. Since this error is very small compared with the standard error of the mean slope, 4·95 ± 0·50, the random experimental error in the relative volumes over this temperature range is almost entirely due to the uncertainty in the slope of the isotherms. At 150° C, however, the difference between the extrapolated value of A calculated from equation (10) and the value obtained from Kleppa's ultrasonic measurements is 0·011; the larger discrepancy

Table 9. *Relative volumes for liquid mercury isotherms and their estimated standard error (computed from Hudleston's relationship for mercury)*

P	*Relative volumes* (V_P^T/V_0^0)										
kg.cm^{-2}	−30° C	−20° C	−10° C	0° C	10° C	20° C	30° C	40° C	50° C	100° C	150° C
0	0·994 56	0·996 37	0·998 19	1·000 00	1·001 82	1·003 63	1·005 45	1·007 27	1·009 10	1·018 25	1·027 48
1000	0·990 94	0·992 70	0·994 46	0·996 22	0·997 99	0·999 76	1·001 52	1·003 29	1·005 06	1·013 9$_3$	1·022 8$_6$
2000		0·989 14	0·990 86	0·992 58	0·994 30	0·996 02	0·997 73	0·999 45	1·001 17	1·009 7$_8$	1·018 4$_5$
3000		0·985 71	0·987 38	0·989 06	0·990 73	0·992 40	0·994 07	0·995 75	0·997 42	1·005 7$_9$	1·014 2$_0$
4000			0·984 02	0·985 65	0·987 28	0·988 91	0·990 54	0·992 17	0·993 79	1·001 9$_4$	1·010 1$_0$
5000			0·980 75	0·982 35	0·983 94	0·985 53	0·987 12	0·988 70	0·990 2$_9$	0·998 2$_2$	1·006 2$_0$
6000				0·979 14	0·980 70	0·982 25	0·983 80	0·985 35	0·986 9$_0$	0·994 6$_3$	1·002 5$_0$
7000				0·976 03	0·977 55	0·979 07	0·980 58	0·982 10	0·983 6$_1$	0·991 1$_5$	0·998 5$_0$
8000					0·974 50	0·975 98	0·977 46	0·978 94	0·980 4$_1$	0·987 7$_8$	0·995$_{00}$
9000					0·971 52	0·972 98	0·974 42	0·975 87	0·977 3$_1$	0·984 5$_1$	0·991$_{50}$
10000						0·970 05	0·971 47	0·972 89	0·974 3$_0$	0·981 3$_3$	0·988$_{50}$
11000						0·967 21	0·968 59	0·969 98	0·971 3$_6$	0·978 2$_5$	0·985$_{00}$
12000							0·965 79	0·967 15	0·968 5$_1$	0·975 2$_5$	0·982$_{00}$
Freezing pressure kg.cm^{-2}	1740	3710	5670	7640	9620	11600					

P	*Standard error of* $V_P^T/V_0^0 \times 10^5$			
kg.cm^{-2}	0° C	50° C	100° C	150° C
2000	± 2	± 2	±11	± 25
4000	8	9	27	50
6000	17	19	43	75
8000	—	31	60	100
10000	—	46	79	125
12000	—	63	96	150

results in greater uncertainty in the relative volumes at 100 and 150° C.

In addition to the experimental errors in the ultrasonic measurements and the P–V–T data, both investigations may be subject to systematic errors. Since there are no other measurements of the compression of mercury under high pressure, only an indirect assessment of such errors is possible. This might be done by comparing the compression of water, as determined by Bridgman[45] simultaneously with that of mercury, with more recent measurements on this liquid. Unfortunately, the compression of water is not known with sufficient accuracy above 1000 atm, and the comparison does not provide a sufficiently direct and accurate estimate of the systematic error for mercury.

Until the velocity of sound in mercury is measured at temperatures below 0° C, some uncertainty in the figures below this temperature must exist. However, since equation (10) represents the intercept as a function of temperature to within 0·001 % over the range 0 to 70° C, there is some justification for extrapolating it to the freezing point of mercury. The difference between the relative volumes given in Table 9 and those published by Bridgman[35] is larger than the estimated random experimental error at pressures below about 4000 kg/cm². Although certain values differ from Bridgman's by as much as 0·00085, we believe our correlation of his results with data obtained from the velocity of sound determinations to be the best that is possible with the existing measurements.

CALCULATION OF VARIOUS THERMODYNAMIC FUNCTIONS FROM THE PARAMETERS OF HUDLESTON'S EQUATION

According to Hudleston's law of intermolecular force at any temperature T° C,

$$L^2P = \lambda(L_0 - L)\exp\left[\theta(L_0 - L)\right] \tag{11}$$

where $$L = \left(\frac{V_P^T}{V_0^t}\right)^{\frac{1}{3}} \text{ and } L_0 = \left(\frac{V_0^T}{V_0^t}\right)^{\frac{1}{3}}$$

The intercept A and the slope B of the Hudleston isotherms are related to the constants λ and θ;

$$A = \log_{10}\lambda \quad \text{and} \quad B = \theta\log_{10}e$$

Differentiating L with respect to P and noting that L_0 and θ are independent of pressure, we obtain

$$\left(\frac{dL}{dP}\right)_T = -\frac{L(L_0 - L)}{P[2(L_0 - L) + L\theta(L_0 - L) + L]} \tag{12}$$

Since $$\left(\frac{dL}{dP}\right)_T = \frac{1}{3L^2V_0^t}\left(\frac{dV}{dP}\right)_T \tag{13}$$

Combining equations (12) and (13), it can be shown that

$$\left(\frac{dV}{dP}\right)_T = -\frac{3L^3V_0^t(L_0 - L)}{P[2(L_0 - L) + L\theta(L_0 - L) + L]} \tag{14}$$

Now, the isothermal coefficient of compressibility at a temperature T is given by

$$\beta^T = -\frac{1}{V_0^T}\left(\frac{dV}{dP}\right)_T \tag{15}$$

Hence $$\beta^T = -\frac{3L^3(L_0 - L)}{PL_0^3[2(L_0 - L) + L\theta(L_0 - L) + L]} \tag{16}$$

Differentiating the temperature dependent terms, L, L_0 and λ in equation (11), with respect to T and noting that the slope of the isotherms $\theta\log_{10}e$ is independent of temperature, we obtain

$$\left(\frac{dL}{dT}\right)_P = \frac{L}{2\lambda}\left(\frac{d\lambda}{dT}\right)_{P=0} + \frac{L}{2(L_0 - L)}\left[\left(\frac{dL}{dT}\right)_{P=0} - \left(\frac{dL}{dT}\right)_P\right]\left[1 + \theta(L_0 - L)\right] \tag{17}$$

Differentiating L_0 and L, with respect to T,

$$\frac{dL_0}{dT} - \frac{dL}{dT} = \frac{1}{3V_0^t}\left[\frac{1}{L_0^2}\left(\frac{dV}{dT}\right)_{P=0} - \frac{1}{L^2}\left(\frac{dV}{dT}\right)_P\right] \tag{18}$$

Combining equations (17) and (18), we obtain that

$$\left(\frac{dV}{dT}\right)_P = \frac{\dfrac{3L^3V_0^t}{2\lambda}\left(\dfrac{d\lambda}{dT}\right)_{P=0} + \dfrac{L^3}{2(L_0)^2}\left(\dfrac{dV}{dT}\right)_{P=0}\left(\dfrac{1}{L_0 - L} + \theta\right)}{1 + \dfrac{L}{2}\left(\dfrac{1}{L_0 - L} + \theta\right)} \tag{19}$$

For mercury and most other liquids except water, the intercept A is a linear function of temperature and can be represented by

$$A^T = \log_{10}\lambda^T = A^\circ - CT$$

Thus $$\frac{1}{\lambda}\frac{d\lambda}{dT} = \frac{-C}{\log_{10}e} = -2{\cdot}3026\,C$$

Hence

$$\left(\frac{dV}{dT}\right)_P = \frac{\dfrac{L^3}{2(L_0^2)}\left(\dfrac{dV}{dT}\right)_{P=0}\left(\dfrac{1}{L_0 - L} + \theta\right) - 3{\cdot}4539L^3V_0^tC}{1 + \dfrac{L}{2}\left(\dfrac{1}{L_0 - L} + \theta\right)} \tag{20}$$

The numerical evaluation of the derivatives $\left(\dfrac{dV}{dT}\right)_P$ and $\left(\dfrac{dV}{dP}\right)_T$ is comparatively easy, as V_0^t, L_0 and $\left(\dfrac{dV}{dT}\right)_{P=0}$ can

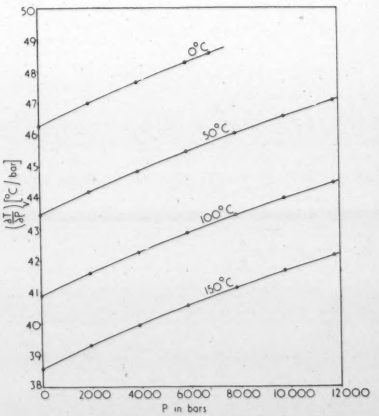


Fig. 6. The pressure-temperature coefficients of mercury as a function of pressure at various temperatures

be calculated from thermal expansion data at atmospheric pressure, while L, θ ($= 2{\cdot}3026\ B$) and C are obtained from the Hudleston relationships for the liquid.

Having evaluated these derivatives, we are able to calculate a number of thermodynamic properties, such as the coefficient $\left(\frac{dE}{dV}\right)_T$, the difference between the heat capacity at constant pressure and constant volume, and the pressure-temperature coefficients $\left(\frac{dP}{dT}\right)_V$.

The pressure-temperature coefficient is given by

$$\left(\frac{dP}{dT}\right)_V = -\left(\frac{dV}{dT}\right)_P \Big/ \left(\frac{dV}{dP}\right)_T$$

Hence, by combining equations (14) and (20),

$$\left(\frac{dP}{dT}\right)_V = \frac{P}{3L_0^2V_0^1}\left(\frac{dV}{dT}\right)_{P=0}\left(\frac{1}{L_0 - L} + \theta\right) - 2{\cdot}3026CP \quad (21)$$

This equation has been solved for liquid mercury and values of the coefficient have been plotted against the pressure for various temperatures in Fig. 6.

CONCLUSION

The preceding discussion of Hudleston's equation has sought chiefly to justify its use in the re-evaluation of the isothermal compressions of mercury. Mercury isotherms of high accuracy are indispensable in many exact measurements at high pressures and our revised figures will be used in the operation of a new primary pressure standard. The standard errors of the mercury compressions have been estimated as precisely as possible, but a final analysis to ensure that the existing data are not subject to any systematic errors will not be possible until new experimental work, perhaps including sound velocity measurements under pressure, has been performed.

The equation accurately represents the best available measurements of liquid compressions, and the properties of the two characteristic parameters permit a straightforward correlation of the data. This will greatly reduce the experimental work needed to improve the extensive early compression results of Amagat, Bridgman and others.

The empirical success of Hudleston's equation has obvious relevance to the development of the theory of the liquid state, and raises other questions (such as the way in which the parameters vary from liquid to liquid), which are not considered here.

ACKNOWLEDGEMENTS

One of us (K. E. B.) would like to express his gratitude to the British Electrical and Allied Industries Research Association for provision of financial assistance during the early stages of this work.

REFERENCES

(1) TUMLIRZ, O. *S.B. Osterr. Akad. Wiss.*, **118**, IIA, p. 203 (1909).
(2) MACLEOD, D. B. *Trans. Faraday Soc.*, **33**, p. 694 (1937).
(3) BRINKMAN, H. C. *Physica*, **7**, p. 747 (1940).
(4) TAIT, P. G. *Voy. Challenger Rep.* Vol. 2 (Physics and Chemistry) (IV), Report on some of the properties of fresh and sea water (London: H.M. Stationery Office, 1888).
(5) GIBSON, R. E. *J. Amer. Chem. Soc.*, **56**, p 4 (1934)
(6) GIBSON, R. E., and KINCAID, J. F. *J. Amer. Chem. Soc.*, **60**, p. 511 (1938).
(7) ADAMS, L. H. *J. Amer. Chem. Soc.*, **53**, p. 3769 (1931).
(8) BRIDGMAN, P. W. *J. Chem. Phys.*, **3**, p. 597 (1935).
(9) RICHARDSON, J. M., ARONS, A. B., and HALVERSON, R. R. *J. Chem. Phys.*, **15**, p. 785 (1947).
(10) HUDLESTON, L. J. *Trans Faraday Soc.*, **33**, p. 97 (1937).
(11) BRIDGMAN, P. W. *International Critical Tables*, **3**, p. 40 (1926).
12) BRIDGMAN, P. W. *Proc. Amer. Acad. Arts. Sci.*, **49**, p. 3 (1913).
(13) AMAGAT, E. H. *Ann. Chim. (Phys.)*, **29**, p. 505 (1893).
(14) NEWITT, D. M., and WEALE, K. E. *J. Chem. Soc.*, p. 3092 (1951).
(15) GIBSON, R. E. *J. Amer. Chem. Soc.*, **57**, p. 1551 (1935).
(16) GIBSON, R. E., and LOEFFLER, O. H. *J. Amer. Chem. Soc.*, **61**, p. 2515 (1939).
(17) GIBSON, R. E. *J. Amer. Chem. Soc.*, **63**, p. 898 (1941).
(18) TILTON, L. W., and TAYLOR, J. K. *J. Res. Nat. Bur. Stand.*, **18**, p. 205 (1937).
(19) TIMMERMANS, J. *Physico-chemical constants of pure organic compounds* (New York: Elsevier Pub. Co. Inc., 1950).
(20) TYRER, D. *J. Chem. Soc.*, **103**, p. 1675 (1913).
(21) TYRER, D. *J. Chem. Soc.*, **105**, p. 2534 (1914).
(22) VILLEY, J. *C.R. Acad. Sci.* [*Paris*], **206**, p. 655 (1938).
(23) PHILIP, N. M. *Proc. Indian Acad. Sci.*, **9A**, p. 109 (1939).
(24) FREYER, E. B., HUBBARD, J. C., and ANDREWS, D. H. *J. Amer. Chem. Soc.*, **51**, p. 759 (1929).
(25) LAGEMANN, R. T., MCMILLAN, D. R., and WOOLF, W. E. *J. Chem. Phys.*, **17**, p. 369 (1949).
(26) COHEN, E., and BUIJ, J. S. *Z. Phys. Chem., B*, **35**, p. 270 (1937).
(27) PATTERSON, T. S. *J. Chem. Soc.*, **81**, p. 1097 (1902).
(28) YOUNG, S. *Sci. Proc. Roy. Dublin Soc.*, **12**, p. 374 (1910).
(29) TRÉHIN, R. *Ann. Phys.* [*Paris*] (IX), **15**, p. 246 (1921).
(30) HUFFMAN, H. M., PARKS, G. S., and DANIELS, A. C. *J. Amer. Chem. Soc.*, **52**, p. 1547 (1930).
(31) STULL, D. R. *J. Amer. Chem. Soc.*, **59**, p. 2726 (1937).
(32) OLIVER, G. D., EATON, M., and HUFFMAN, H. M. *J. Amer. Chem. Soc.*, **70**, p. 1502 (1948).
(33) SEITZ, W., and LECHNER, G. *Ann. Phys.* [*Leipzig*], **49**, p. 93 (1916).
(34) BRIDGMAN, P. W. *Proc. Amer. Acad. Arts. Sci.*, **74**, p. 399 (1942).
(35) BRIDGMAN, P. W. *Proc. Amer. Acad. Arts. Sci.*, **47**, p. 347 (1911).
(36) HARLOW, F. J. *Proc. Phys. Soc.* [*London*], **26**, p. 96 (1913).
(37) HUBBARD, J. C., and LOOMIS, A. L. *Phil. Mag.*, **5**, p. 1177 (1928).
(38) KLEPPA, O. J. *J. Chem. Phys.*, **17**, p. 668 (1949).
(39) RINGO, G. R., FITZGERALD, J. W., and HURDLE, B. G. *Phys. Rev.*, **72**, p. 87 (1947).
(40) HARLOW, F. J. *Phil. Mag.*, **7**, p. 674 (1929).
(41) BEATTIE, J. A., BLAISDELL, B. E., KAYE, J., GERRY, H. T., and JOHNSON, C. A. *Proc. Amer. Acad. Arts. Sci.*, **74**, p. 371 (1941).
(42) DOUGLAS, T. B., BALL, A. F., and GINNINGS, D. C. *J. Res. Nat. Bur. Stand.*, **46**, p. 334 (1951).
(43) BUSHEY, R. H., and GIAUQUE, W. F. *J. Amer. Chem. Soc.*, **75**, p. 806 (1953).
(44) MICHELS, A., WASSENAAR, T., and BLAISSE, B. *Physica*, **9**, p. 574 (1942).
(45) BRIDGMAN, P. W. *Proc. Amer. Acad. Arts. Sci.*, **47**, p. 441 (1912).

On the application of Ångström's method of measuring thermal conductivity

By C. H. Bosanquet, M.A., and R. Aris, B.Sc., Imperial Chemical Industries Ltd., Billingham, Co. Durham

[*Paper received* 6 *October*, 1953]

Ångström's well-known method of measuring thermal conductivity by the steady periodic motion of heat along a bar has a very elegant and simple theory. Its application, however, suffers from the fact that the theory makes use of the first Fourier component of the heat wave. By standardizing the heat input to a rectangular waveform and using the method of calculation here described, accurate results can be obtained without having recourse to a Fourier analysis.

In Ångström's method, a bar of the material, the conductivity of which is to be measured, is heated periodically at one end, the sides and the other end being allowed to radiate to the atmosphere. After a few heating cycles the temperature at any point becomes a periodic function of the same frequency as that of the heat input. If the bar is sufficiently thin for the flow of heat to be linear, Ångström[1] showed that the conductivity of the bar is given by a strikingly simple formula, which is independent of the nature of the bar's surface. The drawback to his method is that this formula [see equation (7) below] contains the amplitude ratio and phase difference of the first Fourier components of the temperature waves at two points. It is neither easy to heat the bar with a sinusoidal input, thereby eliminating all higher Fourier components than the first, nor to perform a complete Fourier analysis of the temperature curve. Indeed, such a Fourier analysis is liable to be inaccurate, since it involves measuring temperatures at at least twelve equally-spaced points of a cycle.

The method, however, is extremely convenient; it can be fitted with automatically recording instruments and left to run overnight and, as has been recently shown,[2] is very suitable for use with poor conductors. The following method of calculation was designed to retain a simple heating cycle, namely that produced by having a heater switched on for half a cycle and off for half a cycle, and to obviate the necessity for a Fourier analysis. Measurements of temperature and time are made at the points where this can be most accurately done; temperatures are measured at the extrema and time is measured at the points where the curves cut the mean temperature line. For convenience the points at which the temperature is measured will be referred to as the thermocouples.

THE THEORY OF THE METHOD

The constants of the material and dimensions of the specimen are:

K, the thermal conductivity,
ρ, the density,
c, the specific heat of the material,
H, the coefficient of emission of its surface,
p, the perimeter,
q, the cross-sectional area of the bar.

Then $D = K/\rho c$ is known as the thermal diffusivity of the material and $\epsilon = Hp/\rho cq$ as the emissivity of its surface. In what follows only the two last constants will occur; in C.G.S. units these are cm^2/sec and sec^{-1} respectively.

If $\theta(x, t)$ is the temperature of the bar above its surroundings at time t and distance x from the heated end, the equation for the linear flow of heat is

$$\frac{\partial\theta}{\partial t} = D\frac{\partial^2\theta}{\partial x^2} - \epsilon\theta \tag{1}$$

If it is assumed, as is the case in practice, that the end distant from the heater is at the temperature of the surroundings, the bar is effectively semi-infinite.

It is convenient to make this equation dimensionless by introducing the variables:

$\xi = x/2\sqrt{(DT)}$, the reduced distance,
$\tau = t/T$, the reduced time,
$\eta = \sqrt{(\epsilon T)}$, the reduced emissivity,
where T = the periodic time of the heating cycle.

Equation (1) now becomes

$$\frac{\partial\theta}{\partial\tau} = \frac{1}{4}\frac{\partial^2\theta}{\partial\xi^2} - \eta^2\theta \tag{2}$$

It may be readily shown that the steady state solution in which the temperature profile at any point is constant in time is,

$$\theta = \sum_{n=0}^{\infty} A_n \exp(-\gamma_n\xi)\cos(2\pi n\tau - \gamma'_n\xi + \phi_n) \tag{3}$$

where the coefficients An are chosen to fit the heat input and

$$\gamma_n, \gamma'_n = \sqrt{\left\{2\left[\pm\eta^2 + \sqrt{(\eta^4 + 4\pi^2n^2)}\right]\right\}} \tag{4}$$

If A'_1 and A'_2 are the amplitudes of the first components at the two points $x = x_1$ and $x = x_2$, and B is the phase difference between the two first components, then

$$\ln A'_1/A'_2 = \gamma_1 l/2\sqrt{(DT)} \tag{5}$$

$$B = \gamma'_1 l\sqrt{(T)}/4\pi\sqrt{(D)}\ (\text{sec}) \tag{6}$$

where $l = x_2 - x_1$ (cm)

But by equation (4) $\gamma_1\gamma'_1 = 4\pi$, hence by equations (5) and (6)

$$D = \tfrac{1}{2}l^2/[\ln(A'_1/A'_2)\,.\,B]\ (\text{cm}^2/\text{sec}) \tag{7}$$

an expression which is independent of ϵ.

From graphs of the first Fourier components A'_1, A'_2 and B could be measured directly. If, however, the actual temperatures are taken and the phase difference estimated from the time lag between two corresponding points, for example, the maxima, then the formula (7) can be very considerably in error.

By standardizing the heat input it is possible to calculate the ratio of the actual temperature amplitude to the amplitude of the first Fourier component, and the difference between

the times at which the actual wave and the first component cross the mean line.

The simple input corresponding to a heater switched on for $\frac{1}{2}T$ sec and off for the following $\frac{1}{2}T$ sec is

$$\left.\begin{aligned} \partial\theta/\partial\xi &= -C(\text{constant}) \quad && m < \tau \leqslant m + \tfrac{1}{2} \\ &= 0 && m + \tfrac{1}{2} < \tau \leqslant m + 1 \\ m &= 0, 1, 2, \ldots \end{aligned}\right\} \quad (8)$$

For this the Fourier series (3) gives the temperature if

$$\left.\begin{aligned} A_0 &= -C/4\eta, \quad \phi_0 = 0 \\ A_n &= 0, \; n \text{ even} \\ &= -2C/\pi n\sqrt{(\gamma_n^2 + \gamma_n'^2)}, \; n \text{ odd} \\ \phi_n &= \tan^{-1}(\gamma_n/\gamma_n') \end{aligned}\right\} \quad (9)$$

In the limiting case of a perfectly lagged bar, $\eta = 0$, and the cycle of heating and not heating must be replaced by one of heating and cooling to avoid an infinite mean temperature. In this case

$$\left.\begin{aligned} A_n &= 0 \quad n \text{ even} \\ &= -2C/(2\pi n)^{\frac{1}{2}}, \quad \phi_n = \tfrac{1}{4}\pi \end{aligned}\right\} \quad (10)$$

The Fourier series is not very convergent, particularly when η is not small, and it is worthwhile obtaining another form of solution as follows. The heating cycle may be thought of as the superposition of a steady heating beginning at $\tau = 0$, a steady cooling of the same intensity beginning at $\tau = \frac{1}{2}$, a further heating beginning at $\tau = 1$ and so on. Thus if $\Theta(\xi, \tau)$ is the solution of equation (2) under steady heating conditions, i.e.

$$\partial\theta/\partial\xi = -C, \quad \xi = 0, \quad \tau \geqslant 0 \quad (11)$$

the solution for the standard heating cycle will be

$$\theta(\tau, \xi) = \sum_{n=0}^{\infty}(-1)^{n+1}\Theta(\tau + \tfrac{1}{2}n, \xi) \quad (12)$$

The solution of equation (2) under the condition (11) may be obtained easily by the use of the Laplace transform. It is

$$\Theta(\tau, \xi) = \frac{C}{2\sqrt{\pi}}\int_0^{\tau} \exp -\left(\eta^2\lambda + \frac{\xi^2}{\lambda}\right)\frac{d\lambda}{\sqrt{\lambda}} \quad (13)$$

which after a little manipulation gives

$$4\eta\Theta/C = \exp(-2\eta\xi) \,.\, [1 + (sgn X)\,\mathrm{erf}\,|X|] - \exp(2\eta\xi) \,.\, \mathrm{erfc}\, Z \quad (14)$$

$$Z, X = \eta\sqrt{\tau} \pm \xi/\sqrt{\tau}$$

where $sgn X$ denotes the sign of X and

$$\mathrm{erf}\, x = 1 - \mathrm{erfc}\, x = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt$$

The series (12) is most convergent when η is large so that the two solutions (12), (14) and equations (3), (9) or (10) provide useful formulae over the whole range of η. Calculations of the amplitude and time of mean were performed for selected values of ξ and η in the range $0 \leqslant \xi \leqslant 1$, $0 \leqslant \eta \leqslant 3$. The positions of the thermocouples, x, and the periodic time T are at the experimeter's disposal to bring the values of ξ and η within this range.

THE METHOD OF CALCULATION

In setting up equation (1) it was assumed that the specimen was a thin rod, so that the temperature was sensibly constant over any cross-section, and that D and ϵ were constant. The errors introduced if the first assumption is wrong will be discussed later, the effect of D varying with temperature must be briefly considered here as it affects the handling of the experimental results. A full discussion of Ångström's method when D and ϵ are not constant has been given by Hagström.[3]

If D and ϵ depend on temperature, the symmetry of the temperature curve about its mean line will be destroyed by the introduction of terms of even order into the Fourier series. The line drawn to be equidistant from the lines of the extrema is thus displaced from the line drawn so as to cut the curve at equal intervals. Fig. 1 illustrates this from

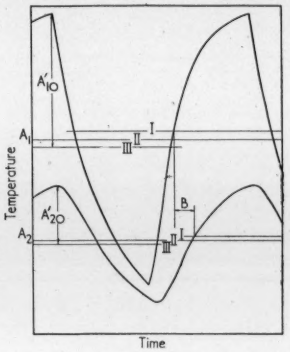


Fig. 1. Typical temperature waveforms

actual experimental data, the difference between lines I and III in the upper curve being some 20° C. To the same order of approximation the true mean temperature line is II, the line half-way between I and III.

From the temperature records taken at two points, x_1 and x_2, the actual amplitude of the temperature variation may be found from the mean of the differences between at least three successive maxima and minima. The line of equal intervals, I, and the line midway between the extrema, III, may be drawn by inspection and the true mean line, II, taken half-way between these. The time lag between the points at which the curves cross their lines of equal interval, I, may then be measured, the mean of several measurements being taken. This is shown in Fig. 1.

Let the mean temperatures of the two waves be A_1 and A_2, their actual amplitudes A'_{10} and A'_{20}, and the measured time lag B_0. Equations (3) and (4) show that the mean temperature at any point is proportional to $\exp(-2\xi\eta)$. It follows that

$$\epsilon = \kappa^2 D \quad (15)$$

where

$$\kappa^2 l = \ln(A_1/A_2)$$

From the actual measurements, a first approximation to D can be obtained by using Ångström's formula (7), namely

$$D_0 = \tfrac{1}{2}l^2/\ln(A'_{10}/A'_{20}) \,.\, B_0 \quad (16)$$

From this the quantities $\eta_0 = \kappa\sqrt{(D_0 T)}$, $\xi_{10} = x_1/2\sqrt{(D_0 T)}$ and $\xi_{20} = x_2/2\sqrt{(D_0 T)}$ can be calculated and the first corrections obtained from Figs. 2 and 3. α_{10} is the value of α in Fig. 2 at the point (ξ_{10}, η_0); it should be noted that the

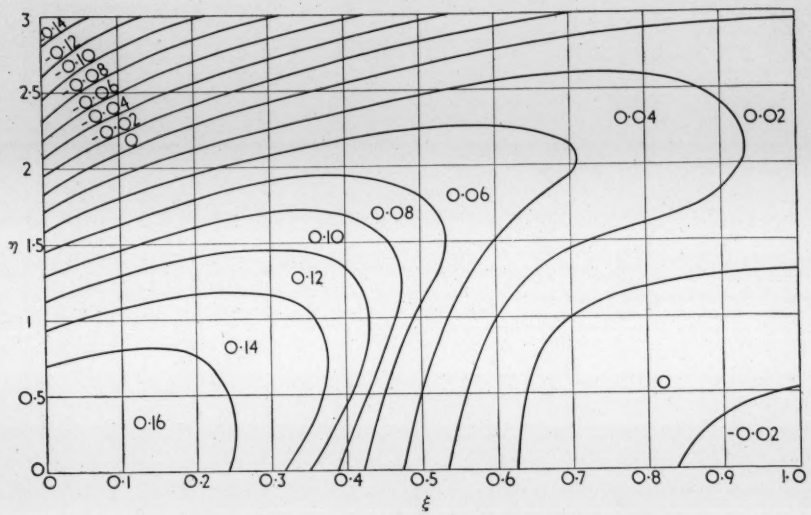


Fig. 2. Chart of $\alpha(\xi, \eta) = \ln$ (actual amplitude/amplitude of first component)

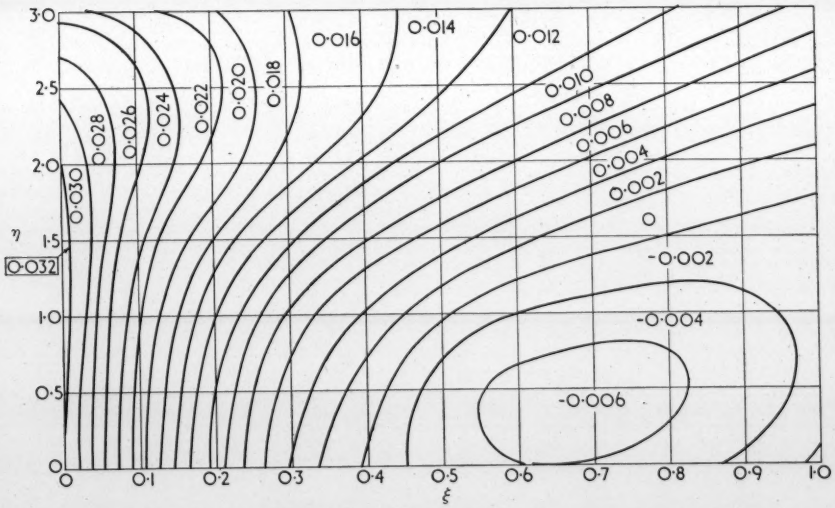


Fig. 3. Chart of $\beta(\xi, \eta) =$ (time of mean of first component − time of mean of actual wave)/(period of heating cycle)

zones of Fig. 2 are labelled with the mean values of α so that, if a point falls on or very close to a contour line, α should be taken as the mean of the values on either side. α_{20}, β_{10} and β_{20} are found similarly and applied to give corrected values

$$\ln (A'_{11}/A'_{21}) = \ln (A'_{10}/A'_{20}) - (\alpha_{10} - \alpha_{20}) \quad (17)$$

$$B_1 = B_0 - T(\beta_{10} - \beta_{20}) \quad (18)$$

Ångström's formula is again used to give a corrected value

$$D_1 = \tfrac{1}{2}l^2/\ln (A'_{11}/A'_{21}) \,.\, B_1 \quad (19)$$

and the correction cycle repeated.

This process can be performed very rapidly and it will often be found that D_1 and D_2 are not sensibly different. If they are

$$D = D_2 - (D_1 - D_2) \Big/ \left(\frac{D_0 - D_1}{D_1 - D_2} - 1\right) \quad (20)$$

is a much better approximation which can be used to start a new correcting cycle.

It should be particularly noted that all corrections are applied to the original values of $\ln (A'_{10}/A'_{20})$ and B_0. Thus

$$\ln (A'_{1n}/A'_{2n}) = \ln (A'_{10}/A'_{20}) - (\alpha_{1n} - \alpha_{2n}) \quad (21)$$

$$B_n = B_0 - T(\beta_{1n} - \beta_{2n}) \quad (22)$$

$$D_n = \tfrac{1}{2}l^2/\ln (A'_{1n}/A'_{2n}) \,.\, B_n \quad (23)$$

defines a sequence D_n which rapidly converges to D.

THE ACCURACY OF THE METHOD

Ångström's formula shows that inaccuracy in the measurement of l is particularly dangerous, and the greatest possible care should be taken over this. Any device which increases the accuracy of the temperature measurement is beneficial and as has been pointed out this method of handling the results makes the most accurate possible use of the record. The difficulty of dividing the cycle into twelve equal parts and the inaccuracy of measurement of these twelve temperatures makes a complete Fourier analysis less accurate as well as more tedious than this method.

A few calculations have been performed to find the order of error introduced if the specimen is a thick cylinder rather than a thin rod. If R is the radius of the cylinder and temperatures are measured on the axis at distances $x_1 = \frac{1}{2}R$, $x_2 = R$ from the heated face, the value of D is found to be approximately $+5\%$ in error. With $x_1 = \frac{3}{2}R$, $x_2 = 2R$ the error drops to less than 1%. It is suggested therefore that when it is not practicable to use a thin specimen the nearer thermocouple should not be at distance from the heated end less than the radius of gyration of the cross-section.

REFERENCES

(1) ÅNGSTRÖM, A. J. *Phil. Mag.*, **25**, p. 130 (1863).
(2) BILLINGTON, N. S. *J. Sci. Instrum.*, **26**, p. 20 (1949).
(3) HAGSTRÖM, K. L. *Ofvers. Vetensk Akad.Förh., Stockh.*, **48**, p. 45 (1891).

The anode fall in a glow discharge

By C. J. F. CHAUNDY, M.A.,* The Clarendon Laboratory, Oxford

[*Paper first received* 29 *December*, 1953, *and in final form* 11 *January*, 1954]

Observations of the electric field in the anode fall region of a glow discharge in nitrogen are made here by measuring the deflexion of an electron beam. The results with a plane anode show agreement with theory: the anode fall is approximately the same as the ionization potential of the gas, and the width of the anode dark space is of the size expected.

PREVIOUS WORK

Measurements of the anode fall of potential have been made by means of solid probes in the vicinity of the anode.(1) As this region is not a quasi-neutral plasma the interpretation of the results is doubtful. Probes and visual methods have been used to measure the width of the anode dark space.(2) For large fields Brose used the Stark-effect broadening of spectral lines to measure the field strength in the cathode dark space of a glow discharge while Aston(3) measured the field by the deflexion of a beam of electrons passing through the discharge. A summary of the measurements and the theory of the anode fall is to be found in von Engel's paper.(4)

APPARATUS

In principle the method used is the same as Aston's. The apparatus is a modified version of that used by Little and von Engel (Fig. 1) and consists of two parts, the cylindrical glass discharge chamber D and the electron gun chamber E separated by the copper disk d. A pressure difference is maintained between D and E by continuous pumping: pure nitrogen from a cylinder flows through a needle valve controlling the rate of flow into D at g_1, and through the capillary tube c mounted on d into the low-pressure chamber E. It is pumped off at g_2.

The electron gun (following Bricka and Bruck(5)) incorporates C as the support for the small hemispherical anode a. Concentric with a is the concave shield s, which is held a few hundred volts negative to the filament f. A potential of several kilovolts is maintained between f and a. Electrons from f are focused by the converging field between s and a to form a beam which crosses D and strikes the fluorescent screen S. The deflexion of the beam due to the field near the discharge anode A is measured on S as a function of the distance between A and the electron beam. A is supported from a phosphor-bronze bellows held by a calibrated screw cap providing a vacuum-tight movable joint. The cathode C of the discharge is so far away that a positive column and a true anode fall region exist. All parts are joined by waxed flanges. The electrodes are of aluminium, and their backs, together with the rods supporting them, are shielded by glass

* Now at Atomic Energy Research Establishment, Harwell.

to prevent unwanted discharges. The gun anode and the discharge anode are both at earth potential.

It was found possible to obtain a spot up to 0·05 mm of mercury with an accelerating voltage on the gun of 3–10 kV and a beam length in the discharge chamber of about 30 cm.

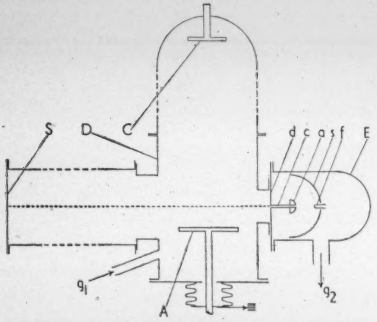


Fig. 1. The discharge chamber and electron gun

Path of electron beam in discharge chamber	= 30 cm
Anode–cathode distance	≃ 25 cm
Diameter of cylindrical discharge chamber	= 6 cm
Diameter of discharge anode *A*	= 5 cm
Length of capillary tube *c*	= 2 cm
Bore of *c*	= 0.2 mm
Diameter of copper anode *a*	= 0.3 cm
Diameter of copper shield *s*	= 3 cm

The dotted line shows the path of the undeflected beam.

The discharge currents were limited by an external resistance of 100 MΩ to several μA only, so that a slightly abnormal glow discharge was maintained.

OBSERVATIONS AND DISCUSSION

It is assumed that the field acting on the electron beam is constant while it is crossing the discharge, and zero outside this region. In fact, then, the beam is deflected towards the anode it moves into stronger fields; however, the deflexion in the discharge is so small (0·01 cm approx.) that this error is negligible.

From the measurements the field X at a distance x from the anode is found. Fig. 2 shows the reduced electric field X/p as a function of px, since the similarity laws are expected to apply to this part of the discharge.[4] Two graphs are given, for two values of the reduced current density j/p^2: the numerical data are given in the figure.

From the figure it is seen that in both graphs the anode fall of potential is approximately 15 V and that the field strength (which here is of the order of 50 V/cm at the most) decreases from the anode monotonically. The reduced width pd of the anode fall region does not alter when the current density is changed, but the field strengths, and therefore the anode fall, increase slightly with increasing j/p^2. It is to be expected that an increase in the pressure p will cause a reduction in d so that pd remains constant for all pressures.

The theory of von Engel indicates that the anode fall should be approximately the ionization potential of the gas (15 V) and that it should increase only very slowly with j/p^2. The measurements here agree with these predictions, and also indicate that the value of pd is about that expected (approximately 0·07 mm of mercury × cm). The predicted dependence of pd on j/p^2 does not, however, appear from the observations.

At higher discharge currents the disk d (see Fig. 1) carried some of the discharge current, for d is at the same potential as the discharge anode A and can act as a subsidiary anode. Under these conditions spurious reversed deflexions were found because the beam passes through a region where the

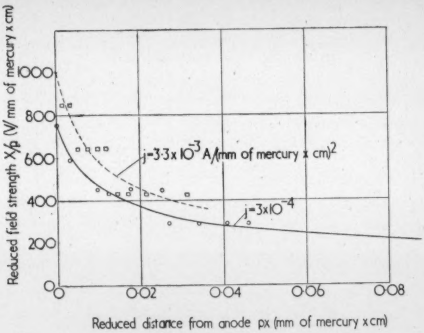


Fig. 2. The reduced field strength as a function of the relative distance from the anode in nitrogen with aluminium electrodes

For full line graph:	
pressure	= 0.05 mm of mercury
discharge voltage	= 600 V
current density	= 3.10^{-4} A/(mm of mercury × cm)^2
accelerating voltage on gun	= 8.9 kV
For dashed graph:	
pressure	= 0.016 mm of mercury
discharge voltage	= 950 V
current density	= 3.3×10^{-3} A/cm^2 mm^2
accelerating voltage on gun	= 4.1 kV

field is distorted. With an improved design it should be possible to make more accurate measurements of the anode fields, for the electron beam does not disturb the discharge. The observations reported here show that the method is reliable and a practical tool for future research which may enable yet smaller fields to be measured.

ACKNOWLEDGEMENTS

I wish to thank Mr. T. C. Keeley for extending to me the facilities of the Clarendon Laboratory, Dr. A. von Engel for suggesting the problem and Dr. P. F. Little for guidance and help in preparing the paper.

REFERENCES

(1) Guntherschulze, A., and Schnitger, H. *Z. Phys.*, **104**, p. 395 (1936).
(2) Brose, E. *Ann. Phys.* [*Leipzig*], **58**, p. 731 (1919).
(3) Aston, F. W. *Proc. Roy. Soc. A*, **84**, p. 526 (1911).
(4) von Engel, A. *Phil. Mag.*, **32**, p. 417 (1941).
(5) Bricka, M., and Bruck, H. *Ann. Radioélect.*, **3**, p. 1 (1948).

The determination of grain size from spotty X-ray diffraction rings

By P. B. Hirsch, M.A., Ph.D., A.Inst.P., Crystallographic Laboratory, Cavendish Laboratory, Cambridge

[*Paper received* 30 *July*, 1953]

A study is made of the effects of a distribution of grain size, a large absorption coefficient and overlapping diffraction lines on the values of grain size obtained from the number of spots on X-ray back reflexion photographs from polycrystalline material. It is shown that accurate values of the mean grain size can be obtained by extrapolation to high exposure ratios and that the general "volume" formula relating the number of spots to the penetration of the beam into the specimen is applicable also in cases in which the average penetration of the beam is small compared with the grain diameter.

1. Introduction

The discovery of the existence in crystals of substructures too ill-defined to be observed under the microscope has opened a new field for the application of X-ray methods for the determination of grain size. By using collimators of suitable diameters it is possible to obtain spotty X-ray diffraction rings from specimens with grain sizes in the range of 1–100 μ, which covers the dimensions most frequently encountered in substructures. In these cases the grain size can be determined from the number of spots on the diffraction rings.[1–3] For thick metal specimens the use of the double exposure method due to Stephen and Barnes[2] is most convenient. Generally, the number of spots on a diffraction ring is inversely proportional to the volume of the grains; thus, if the (linear) grain size is halved, the number of spots is increased by a factor of eight, i.e. the appearance of the diffraction pattern changes radically. It is clear, therefore, that the X-ray method is capable of giving very accurate values of linear grain size. The original results of Stephen and Barnes do not give any indication that the method is capable of such precision. In fact, certain precautions are necessary if a high degree of accuracy is to be obtained. The effects of divergence of the beam, of the finite range of reflexion of the crystals and of the preferred orientation in the specimen have been discussed in some detail in previous publications.[3–5] In this paper the effects of a distribution of grain size, a large absorption coefficient and overlapping diffraction lines will be considered. It will be assumed that the reader is familiar with the paper by Hirsch and Kellar[3] whose notation will be followed.

2. Distribution of Grain Volume

Let $f(v)dv$ be the number of grains with volumes between v and $v + dv$, per unit volume of material. Then, in a volume V of material, corresponding to a penetration l, the number of these grains is equal to $Vf(v)dv$ and the number of reflexions N_v is

$$N_v = \tfrac{1}{2}Ap \cos\theta(d\theta + \Delta)lf(v)dv \qquad (1)$$

Only those spots are counted whose blackening is greater than an arbitrary value B_0, which is related to the penetration l_0 by the equation

$$B_0 = vKI_0T_0 \exp\left[-\mu l_0(1 - \sec 2\theta)\right] \qquad (2)$$

It is clear that for constant values of B_0, T_0 and I_0, l_0 is a function of v. This means simply that the larger grains at greater depths produce spots on the film of the same blackening as smaller grains nearer the surface. The smallest grain producing a spot on the film is one at the surface of volume $v_1 = B_0/T_1I_0K$. Since v_1 depends on the exposure T_1, the volumes of the smallest grains counted on photographs of different exposures are not equal.

The total number of spots on one ring on a photograph is equal to

$$N = \tfrac{1}{2}Ap \cos\theta(d\theta + \Delta)\int_{v_1}^{\infty} lf(v)dv$$

Substituting for l

$$N = C\int_{v_1}^{\infty} \log\frac{vKI_0T_1}{B_0}f(v)dv \qquad (3)$$

where

$$C = \frac{Ap \cos\theta(d\theta + \Delta)}{2\mu(1 - \sec 2\theta)}$$

This can be written in the form

$$N = C\int_0^{\infty} \log\frac{\bar{v}KI_0T_1}{B_0}f(v)dv + C\int_0^{\infty} \log\frac{v}{\bar{v}}f(v)dv - C\int_0^{v_1} \log\frac{vKI_0T_1}{B_0}f(v)dv$$

$$= C\log\frac{\bar{v}KI_0T_1}{B_0} + C\int_0^{\infty} \log\frac{v}{\bar{v}}f(v)dv + C\int_0^{v_1} \log\frac{v_1}{v}f(v)dv \qquad (4)$$

where $\bar{v} = 1/\int_0^{\infty} f(v)dv$ = mean volume of grains.

The first term is equal to the number of spots for a specimen of uniform grain size $\bar{v}$; the second term is a constant (positive or negative) depending on the distribution function and the third term is positive and decreases with increasing exposure, i.e. decreasing v_1. The relation between N and log (exposure time T_1), or the equivalent mean penetration l, is shown schematically in Fig. 1(*a*). The straight line is predicted by the simple theory (no distribution), i.e. the first term of equation (4); the curve represents the complete expression. The actual number of spots on the photographs may be greater or smaller than that predicted by the simple theory, but the slope of the curve approaches that for the simple theory for long exposures.

In order to obtain estimates of the grain size, two exposures are necessary from which the difference in the number of spots is determined. In the second, and longer, exposure T_2, the smallest grain counted has a volume $v_2 = B_0/T_2I_0K$, smaller than v_1. For grains with $v \geqslant v_1$ the difference in the number of spots on the two photographs is

$$(N_2 - N_1)_{v \geqslant v_1} = C\log\frac{T_2}{T_1}\int_{v_1}^{\infty} f(v)dv$$

In addition, there are a number of spots counted on the longer exposure with $v_2 < v < v_1$. This number is equal to

$$C\int_{v_2}^{v_1} \log \frac{vKI_0T_2}{B_0} f(v)dv$$

$$= C\int_{v_2}^{v_1} \log \frac{v}{v_2} f(v)dv$$

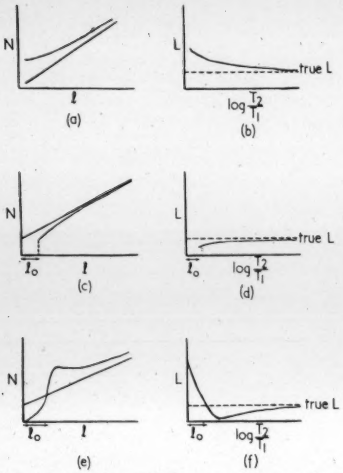


Fig. 1. Schematic variation of the number of spots on the diffraction rings with beam penetration, and of the grain size with log of the exposure ratio, for weakly and strongly absorbing specimens

It follows that the total difference in the number of spots is

$$N_2 - N_1 = C\left[\log \frac{T_2}{T_1}\int_{v_1}^{\infty} f(v)dv + \int_{v_2}^{v_1} \log \frac{v}{v_2} f(v)dv\right] \quad (5)$$

$$= C\left[\log \frac{T_2}{T_1}\int_{v_2}^{\infty} f(v)dv + \int_{v_2}^{v_1} \log \frac{v}{v_1} f(v)dv\right] \quad (6)$$

The second integral is positive in equation (5) and negative in equation (6). Thus we can write

$$N_2 - N_1 = C \log \frac{T_2}{T_1}\int_{v_L}^{\infty} f(v)dv \quad (7)$$

where v_L is a value between v_1, v_2 chosen so that expression (7) is equal to equations (5) and (6). If the distribution function $f(v)$ is negligible for v_1 (and therefore for v_L and v_2 for symmetrical distributions)

$$N_2 - N_1 = C \log \frac{T_2}{T_1}\int_{0}^{\infty} f(v)dv = C \log \frac{T_2}{T_1}\frac{1}{\bar{v}}$$

If $f(v)$ is not negligible for v_L, the grain volume deduced is equal to the mean value of the volume of all grains with $v > v_L$, i.e. the grain volume will always be an over-estimate. In order to obtain a value of the grain volume as near to the true mean as possible, v_1 and v_2 should be as small as possible, i.e. B_0 should be small and T_1, T_2 should be large. If a series of exposures is taken, and the difference in the number of spots is counted relative to an arbitrary blackening on the shortest exposure, the effective grain size should decrease with increasing exposure ratio. This is shown schematically in Fig. 1(*b*); each point on the curve is inversely proportional to the slope of the chord between the two points representing the two exposures on the curve of Fig. 1(*a*). If the exposure ratio is varied, the end of the chord representing the short exposure is fixed, the other end moves along the curve with increasing l, and the slope will consequently increase, reaching the true value predicted by the simple theory asymptotically with large exposure ratios. The physical explanation for the large apparent value of grain size and its variation with exposure ratio, is the fact that for any two exposures the difference in the number of spots counted is always smaller than predicted by the simple theory because the reflexions from the smallest grains are too weak to be observed, but that this effect becomes relatively less important the greater the mean penetration of the beam.

The predicted variation of the apparent grain size is shown very clearly by all the results tabulated by Stephen and Barnes. Some of their figures are reproduced in the table. It is certain that the true grain size is less than 29 μ, and probably very close to 28 μ. It is most important that such systematic trends are recognized in grain size determinations; average values should not be calculated. The mean value estimated by Stephen and Barnes is clearly too high by at least 7%. In practice long exposures should be taken and B_0 reduced as far as possible. It is convenient to take a series of exposures and to extrapolate to infinite exposure ratio. In Fig. 2 the results of Stephen and Barnes are

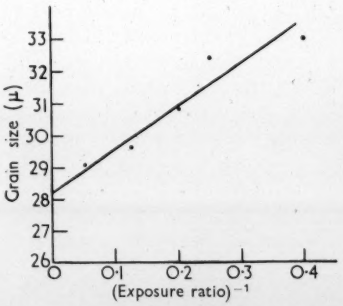


Fig. 2. Extrapolation of Stephen and Barnes' results

plotted against (exposure ratio)$^{-1}$. The extrapolated value is approximately 28·2 μ.

It may be mentioned here that by varying the exposure time, this method might be used to determine the distribution of grain size.

3. THE EFFECT OF LARGE GRAIN SIZE AND ABSORPTION

Stephen and Barnes suggested that if the penetration of the rays into the specimen is small compared with the thickness of the grain, the "volume" formula is no longer applicable. The number of spots is now given by

$$N = \tfrac{1}{2}p\cos\theta\frac{A}{a}(d\theta + \Delta) \qquad (8)$$

where a = area of cross-section of grain. This formula is correct if all the grains extend below the surface by distances large compared with the penetration of the X-ray beam [see Fig. 3(*a*)]. If the surface is a random section through the

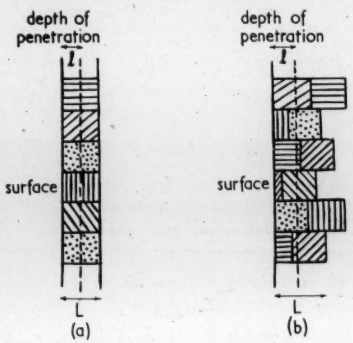


Fig. 3. The relation of depth of penetration to grain size for highly absorbing specimens

(*a*) Equation (8) applicable. (*b*) Equation (9) applicable.

metal, the grains extend below the surface by all possible distances between zero and the length (L) of the grains [Fig. 3(*b*)]. There will be $\frac{A}{a}\times\frac{l}{L}$ grains which terminate between the surface and the depth of penetration l; for each of these, two grains will be irradiated. The total number of grains irradiated is therefore

$$\frac{A}{a}\left(1-\frac{l}{L}\right)+2\frac{A}{a}\frac{l}{L}=\frac{A}{a}\left(1+\frac{l}{L}\right),$$

and hence the correct expression for the number of spots is

$$N = \tfrac{1}{2}p\cos\theta\frac{A}{a}(d\theta + \Delta)\left(1+\frac{l}{L}\right) \qquad (9)$$

This reduces to the surface formula for $\frac{l}{L}\to 0$ and to the volume formula for $\frac{l}{L}\to\infty$. It would appear therefore that the correct grain area can be determined by reducing the exposures sufficiently so that $\frac{l}{L}\ll 1$. In practice, however, there is always a distribution of grain size so that only the reflexions from the largest grains would be counted for the shortest exposures. Further, in order to detect the reflexions at all, the penetration l cannot be reduced below a certain limit, and this limit cannot be calculated. These two effects can lead to values of grain size which are incorrect by an order of magnitude. It is concluded that the "surface" formula should not be applied in practice.

Suppose, however, that two different exposures are taken. Then the difference in the number of spots is

$$N_1 - N_2 = \tfrac{1}{2}p\cos\theta\frac{A}{aL}(l_1 - l_2)$$

$$= \tfrac{1}{2}p\cos\theta\frac{A}{v}(l_1 - l_2)$$

which is identical with the "volume" formula. Thus, even if the penetration into the specimen is small compared with the diameter of the grains, the "volume" formula can be applied in the usual way.

The variation of N with l is a little more complicated in this case. Suppose first that there is no distribution of grain size. Then the theoretical formula, assuming all possible reflexions are counted, extrapolates to $\frac{1}{2}p\cos\theta\frac{A}{v}L(d\theta+\Delta)$ at $l=0$ [Fig. 1(*c*)]. In practice, the spots cannot be detected unless the penetration into the specimen is greater than a limiting value l_0. For any exposure corresponding to a penetration l, there will be some grains whose lengths in the irradiated volume are less than l_0. It is easy to show that the number of spots from grains, whose irradiated lengths are greater than l_0, is equal to $\frac{1}{2}p\cos\theta\frac{A}{v}(L+l-2l_0)$. This expression extrapolates to $\frac{1}{2}p\cos\theta\frac{A}{v}(L-l_0)$ at $l=l_0$; when l increases, l_0 decreases, so that the actual curve approaches the theoretical line [according to equation (9)] for large values of l [see Fig. 1(*c*)]. The slope of this curve decreases with increasing l, and therefore, following the argument used in Section (2), the apparent grain size increases with l. This type of variation is shown in Fig. 1(*d*).

If there is also a distribution of grain size, all the points may be higher or lower, and the resulting curve might, for example, be of the type shown in Fig. 1(*e*). At large values of l, the curve is similar to that in 1(*a*). If a series of exposures is taken and the difference in the number of spots counted relative to that on a short exposure for which the mean penetration is less than l_0, the variation of the difference in the number of spots with exposure ratio can again be obtained by keeping one end of a chord on the curve fixed near the origin, and moving the other point along the curve. The slope of this chord will first increase, then reach a maximum and finally decrease, so that the apparent grain size first decreases, reaches a minimum and then increases, as shown in Fig. 1(*f*). It is now quite clear that unless special precautions are taken it is possible to derive values of apparent grain size which can be considerably larger or smaller than the true value. In order to arrive at an accurate value it is absolutely essential to take long exposures and to extrapolate the results to large values of exposure ratio.

4. THE EFFECT OF OVERLAPPING DIFFRACTION LINES

In some cases several diffraction lines may overlap and corrections become necessary.

(*a*) *Lines of different wavelengths*, e.g. α_1, α_2. The total difference in the number of spots is equal to

$$(N_2 - N_1)_T = (N_2 - N_1)_{\alpha_1} + (N_2 - N_1)_{\alpha_2}$$
$$= C \log \frac{T_2}{T_1}\left[\int_{v_{L_1}}^{\infty} f(v)dv + \int_{v_{L_2}}^{\infty} f(v)dv\right]$$

since the constant C will be very nearly the same for both reflexions. As the intensity of the direct beam for the α_2 line is about half that of the α_1 line, $v_{L_2} \simeq 2v_{L_1}$. If the exposures are sufficiently long so that the distribution function is vanishingly small for v_{L_2},

$$(N_2 - N_1)_T = 2C\frac{1}{\bar{v}} \log \frac{T_2}{T_1},$$

so that the true grain size is equal to $2^{(1/3)} \times$ the apparent grain size. This condition may be satisfied by extrapolating to long exposure times.

(*b*) *Lines of different indices*. If the structure factors for the reflexions are equal, the effect can be allowed for by replacing the multiplicity factor by the sum of the multiplicity factors of the two lines. If the structure factors are not equal but comparable in magnitude, the multiplicity factors should be added and the results extrapolated to infinite exposure times. Finally, if the structure factor for one line is much smaller than that for the other, the contribution of the weak line should be neglected.

Typical results obtained by Stephen and Barnes[2]

Exposure ratio	*Grain size (μ) calculated*		*Metallurgical grain size (μ)*
2½	33		
4	32·4		
5	30·8	mean	
8	29·6	31	28
20	29		

REFERENCES

(1) Schdanow, H. S. *Z. Kryst.*, **90**, p. 82 (1935).
(2) Stephen, R. A., and Barnes, R. J. *J. Inst. Metals*, **60**, p. 285 (1937).
(3) Hirsch, P. B., and Kellar, J. N. *Acta Cryst.*, **5**, p. 162 (1952).
(4) Hirsch, P. B., and Kellar, J. N. *Proc. Phys. Soc. [London] B*, **64**, p. 369 (1951).
(5) Gay, P., and Kelly, A. *Acta Cryst.*, **6**, p. 165 (1953).

The flux linkage with a search coil produced by a coaxial uniformly magnetized prolate spheroid

By G. H. Hunt, B.Sc., Durham Colleges in the University of Durham

[*Paper first received* 7 *August, and in final form* 11 *November*, 1953]

An expression is derived for the flux enclosed by a search coil coaxial with a uniformly magnetized prolate spheroid. The expression takes the form of power series with coefficients which can readily be calculated from mathematical tables.

The intensity of magnetization of a ferromagnetic body is usually measured by using a magnetometer or a search coil and fluxmeter. In the case of a uniformly magnetized ellipsoid of revolution the necessary expressions for the field at points on the axes of symmetry have already been given by Peake and Davy.[1] However, the author is not aware of any expressions for the flux linkage with a multi-layer coil coaxial with an ellipsoid. A general solution of the problem would be difficult to formulate, but the results derived below may be found sufficiently accurate for many practical applications. The only case considered is that of a prolate spheroid, since magnetization in the direction of the axis of revolution of an oblate spheroid is not usually attempted, but the same method could, if necessary, be applied to the latter case.

THEORY

The initial steps in the theory are quoted by Livens[2] and Stratton,[3] and need not be elaborated here. Consider the ellipsoid of revolution

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{b^2} = 1$$

where $b^2 = a^2(1 - e^2)$, $1 > e > 0$

The system of equipotential surfaces outside the ellipsoid is found to coincide with the confocal ellipsoids

$$\frac{x^2}{a^2 + \lambda} + \frac{y^2}{b^2 + \lambda} + \frac{z^2}{b^2 + \lambda} = 1$$

When the applied field H is in the direction of the axis of x, then the potentials inside and outside the ellipsoid are given respectively by

$$\psi_i = \frac{-Hx}{1 + A(\mu - 1)},$$

$$\psi_0 = -Hx + \frac{ab^2}{2}\frac{(\mu - 1)Hx}{1 + A(\mu - 1)}\int_{\lambda}^{\infty} \frac{d\theta}{(a^2 + \theta)^{\frac{3}{2}}(b^2 + \theta)}$$

where μ is the permeability of the medium and

$$A = \frac{ab^2}{2}\int_0^{\infty} \frac{d\theta}{(a^2 + \theta)^{\frac{3}{2}}(b^2 + \theta)}$$

Without any loss of generality z may be made zero and only points in the xy plane need then be considered. The variable which defines the confocal ellipsoid is best replaced by t where

$$t^2 = (a^2 + \lambda)/a^2e^2$$

and hence

$$\frac{x^2}{t^2} + \frac{y^2}{t^2 - 1} = a^2e^2 \qquad (1)$$

Evaluating the integrals we have

$$A = \frac{1 - e^2}{e^2}\left(\frac{1}{e}\coth^{-1}\frac{1}{e} - 1\right) \qquad (2)$$

and

$$\psi_0 = -Hx + \frac{ab^2}{2}\frac{(\mu - 1)Hx}{1 + A(\mu - 1)}\frac{2}{a^3e^3}\left(\coth^{-1} t - \frac{1}{t}\right) \qquad (3)$$

From expressions (2) and (3) the values of the fields parallel to the axis of x may be found by differentiation. Hence the field inside the ellipsoid is

$$H_i = -\frac{\partial \psi_i}{\partial x} = \frac{H}{1 + A(\mu - 1)}$$

But μH_i is the induction of the material, or

$$(\mu - 1)H_i = 4\pi I$$

where I is the intensity of magnetization.

Therefore $$H_i = H - 4\pi IA \qquad (4)$$

$4\pi A$ is therefore the demagnetizing factor of the ellipsoid. Also, the field outside the ellipsoid in the direction of x is given by

$$H_0 = -\frac{\partial \psi_0}{\partial x}$$

$$= H - \frac{(\mu - 1)H}{1 + A(\mu - 1)}\left(\frac{1 - e^2}{e^2}\right)\frac{\partial}{\partial x}\left[\frac{x}{e}\left(\coth^{-1} t - \frac{1}{t}\right)\right]$$

$$= H - 4\pi I\left(\frac{1 - e^2}{e^2}\right)\frac{\partial}{\partial x}\left[\frac{x}{e}\left(\coth^{-1} t - \frac{1}{t}\right)\right]$$

Now $$\frac{\partial}{\partial x}\left[\frac{x}{e}\left(\coth^{-1} t - \frac{1}{t}\right)\right]$$

$$= \frac{1}{e}\left(\coth^{-1} t - \frac{1}{t}\right) + \frac{x}{e}\left(\frac{1}{1 - t^2} + \frac{1}{t^2}\right)\frac{\partial t}{\partial x}$$

and $$x\frac{\partial t}{\partial x} = t(t^2 - 1)\Big/\left(\frac{t^4 a^2 e^2}{x^2} - 1\right)$$

Therefore

$$H_0 = H - 4\pi I\left(\frac{1 - e^2}{e^3}\right)\left[\coth^{-1} t - \frac{1}{t} - \frac{1}{t\left(\frac{t^4 a^2 e^2}{x^2} - 1\right)}\right] \qquad (5)$$

To find the flux enclosed by a loop, the axis of which coincides with the axis of x, only fields parallel to that axis need be considered. The flux required will then be given by

$$\int_{\text{inside}} \mu H_i . 2\pi y dy + \int_{\text{outside}} H_0 . 2\pi y dy$$

But $$y^2 = (t^2 - 1)(a^2 e^2 - x^2/t^2)$$

Therefore $$y dy = (a^2 e^2 t - x^2/t^3)dt \qquad (6)$$

Hence total internal flux

$$= 2\pi a^2 e^2 \int (H_i + 4\pi I) t dt - 2\pi x^2 \int (H_i + 4\pi I) dt/t^3 \qquad (7)$$

and total external flux $$= 2\pi a^2 e^2 \int H_0 t dt - 2\pi x^2 \int H_0 dt/t^3 \qquad (8)$$

At the surface of the ellipsoid $t = 1/e$, and on the axis of x $t = x/ae$. The upper limit for the external flux integral will depend on the size of the loop, and will correspond to some chosen value of t.

Therefore, total internal flux

$$= 2\pi(H + 4\pi I\sqrt{1 - A})\left(a^2 e^2 \int_{x/ae}^{1/e} t dt - x^2 \int_{x/ae}^{1/e} dt/t^3\right)$$

$$= \pi(H + 4\pi I\sqrt{1 - A})(1 - e^2)(a^2 - x^2) \qquad (9)$$

Total external flux

$$= 2\pi H\left(a^2 e^2 \int_{1/e}^{t} t dt - x^2 \int_{1/e}^{t} dt/t^3\right)$$

$$- 8\pi^2 I\left(\frac{1 - e^2}{e^3}\right)\left\{a^2 e^2 \int_{1/e}^{t} t\left(\coth^{-1} t - \frac{1}{t}\right)dt\right.$$

$$- x^2 \int_{1/e}^{t} \frac{1}{t^3}(\coth^{-1} t + 1/t)dt$$

$$\left. - \int_{1/e}^{t} \left(a^2 e^2 t - \frac{x^2}{t^3}\right)\left[\frac{1}{t\left(\frac{t^4 a^2 e^2}{x^2} - 1\right)}\right]dt\right\}$$

$$= \pi a^2\left[He^2 t^2\left(1 + \frac{x^2}{a^2 e^2 t^4}\right)\right.$$

$$\left. - 4\pi I\left(\frac{1 - e^2}{e}\right)\left(1 - \frac{x^2}{a^2 e^2 t^2}\right)\sqrt{(t^2 - 1)}\coth^{-1} t - t)\right]_{1/e}^{t} \qquad (10)$$

Combining expressions (9) and (10), the total flux through the loop is given by

$$F = \pi H\left(a^2 e^2 - \frac{x^2}{t^2}\right)(t^2 - 1)$$

$$- 4\pi^2 a^2 I\left(\frac{1 - e^2}{e}\right)\left(1 - \frac{x^2}{a^2 e^2 t^2}\right)\sqrt{(t^2 - 1}\coth^{-1} t - t) \qquad (11)$$

$$= \pi y^2 H + \pi y^2 . 4\pi I\left(\frac{1 - e^2}{e^3}\right)\left(\frac{t}{t^2 - 1} - \coth^{-1} t\right) \qquad (12)$$

The first part is the flux due to the undisturbed applied field, while the second is proportional to the intensity of magnetization of the specimen. From this it can be seen that the flux per unit area due to the specimen is constant for a fixed value of t, that is to say, for a particular confocal ellipsoid. To determine the flux through a solenoid of finite dimensions coaxial with the ellipsoid, the expression (11) must be integrated with respect to x and y. The integration with respect to x has been carried out by Prache and Cazenave.[4] They find that the flux due to the ellipsoidal specimen through a single layer coil of radius r and n turns per unit length, extending from co-ordinate x_1 to x_2 is

$$F = \pi r^2 n . 4\pi I\left(\frac{1 - e^2}{e^3}\right)$$

$$\left[x\coth^{-1} t + \frac{x}{t}\left(1 + \frac{2a^2 e^2}{3r^2} + \frac{1}{3(t^2 - 1)}\right)\right]_{x_1}^{x_2}$$

and hence for a coil of length $2l$ placed symmetrically about the specimen,

$$F = 8\pi^2 I r^2 ln\left(\frac{1 - e^2}{e^3}\right)\left[\coth^{-1} t + \frac{1}{t} + \frac{2a^2 e^2}{3r^2 t} + \frac{1}{3t(t^2 - 1)}\right] \qquad (13)$$

where $$\frac{l^2}{t^2} + \frac{r^2}{t^2 - 1} = a^2 e^2$$

For a coil of finite thickness, this expression must now be integrated with respect to r, but the author has not found a general solution to the integration. For the particular case of $l = ae$, the integral is readily soluble. It is then found

that for a coil of length $2ae$, internal radius r_1, and external radius r_2, with winding density n_1 and n_2 turns per unit length perpendicular and parallel to its axis, the flux linkage is

$$F = \tfrac{8}{3}\pi^2 Ila^3 n_1 n_2 (1 - e^2)$$

$$\left[\frac{(t^2-1)^3}{t^3}\coth^{-1} t - 2\log_e t + 2t^2 - \frac{1}{2t^4}\right]_{t_1}^{t_2} \quad (14)$$

where $t = ae(t^2 - 1)/t$.

A method of approximation has been developed for the general case, but this is only applicable provided the coil length is less than half the specimen length. Examination of Fig. 1 shows that t varies very little with x until x/ae is greater than $\frac{1}{2}$, hence it is possible to express t fairly accurately as a power series in x/ae.

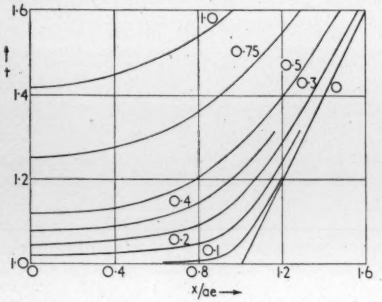


Fig. 1. The variation of t with x/ae. Values of y/ae are shown on the curves

Consider the variation of t with x, and in equation (1) for x against t substitute the values $x = 0$ and $x = x$, corresponding to $t = s$ and $t = s + \delta s$ respectively. Then

$$\left(\frac{x}{ae}\right)^2\left[\frac{(s+\delta s)^2 - 1}{(s+\delta s)^2}\right] = (s+\delta s)^2 - s^2 \quad (15)$$

Ignoring powers of δs greater than the first,

$$\delta s = \frac{(x/ae)^2(s^2-1)}{2s^3 - 2s(x/ae)^2} \quad (16)$$

$$= \frac{(x/ae)^2(s^2-1)}{2s^3}\left[1 + \frac{1}{s^2}(x/ae)^2 + \frac{1}{s^4}(x/ae)^4 \ldots\right] \quad (17)$$

Now consider variation of the term $[t/(t^2-1) - \coth^{-1} t]$ with t. Denoting this by $f(t)$, then by a Taylor expansion, putting $t = (s + \delta s)$

$$f(s+\delta s) = f(s) + \delta s f'(s) + \frac{(\delta s)^2}{2!}f''(s) + \ldots. \quad (18)$$

$$= \frac{s}{s^2-1} - \coth^{-1} s - \frac{2}{(s^2-1)^2}\cdot\delta s + \frac{4s}{(s^2-1)^3}\cdot(\delta s)^2 - \frac{4}{3}\frac{(5s^2+1)}{(s^2-1)}\cdot(\delta s)^3 + \ldots \quad (19)$$

Substituting for δs in terms of x/ae, and ignoring powers of x/ae above the sixth,

$$f(s+\delta s) = \frac{s}{s^2-1} - \coth^{-1} s - \left(\frac{x}{ae}\right)^2\frac{1}{s^3(s^2-1)} + \left(\frac{x}{ae}\right)^6\frac{11s^2+1}{6s^9(s^2-1)} \quad (20)$$

For values of x/ae less than $\frac{1}{2}$, the last term in the expression represents a small correction factor, and since s is very nearly equal to 1, $(11s^2 + 1)$ may be approximated to $12s^2$.

Therefore

$$f(s+\delta s) = \frac{s}{s^2-1} - \coth^{-1} s - \left(\frac{x}{ae}\right)^2\frac{1}{s^3(s^2-1)} + \left(\frac{x}{ae}\right)^6\frac{2}{s^7(s^2-1)} \quad (21)$$

To illustrate the accuracy of this expression, consider the case of $x = ae/2$, $y = ae/4$. Then $s = 1{\cdot}0308$, $s' = 1{\cdot}0399$, and $\delta s = 0{\cdot}0091$.

The expansion for δs gives $\delta s = 0{\cdot}0089$.

The expansion for $f(s + \delta s)$ is composed of the following terms:

$$\frac{s}{s^2-1} - \coth^{-1} s = 14{\cdot}40$$

$$-\left(\frac{x}{ae}\right)^2\frac{1}{s^3(s^2-1)} = -3{\cdot}65$$

$$\left(\frac{x}{ae}\right)^6\frac{2}{s^7(s^2-1)} = 0{\cdot}20$$

Therefore $f(s + \delta s)$ by expansion is $10{\cdot}95$. $f(s')$ is $10{\cdot}83$.

The expansion is therefore accurate to about 1% for the particular values chosen. For smaller values of x/ae the accuracy is higher. Note that, after integration with respect to x, the accuracy will improve, since the higher terms in the series will be reduced compared with the first term.

For x/ae not greater than $\frac{1}{2}$ it is therefore reasonable to take the flux through the loop due to the intensity of magnetization as

$$F = \pi y^2 . 4\pi I\left(\frac{1-e^2}{e^3}\right)\left[\frac{s}{s^2-1} - \coth^{-1} s - \frac{1}{s^3(s^2-1)}\left(\frac{x}{ae}\right)^2 + \frac{2}{s^7(s^2-1)}\left(\frac{x}{ae}\right)^6\right] \quad (22)$$

where s is defined by

$$s^2 - 1 = y^2/a^2e^2$$

The flux due to the applied field need not be further considered since its magnitude for a particular size of search coil is easily calculated.

For a uniformly wound single layer coil of n_1 turns per unit length placed coaxially with the ellipsoid and symmetrically along its length, the total flux-turns through the coil will be

$$G = 2n_1\int_0^l F dx \quad (23)$$

where l is the half-length of the coil.

Therefore,

$$G = 2\pi y^2 n_1 . 4\pi I\left(\frac{1-e^2}{e^3}\right)l\left[\frac{s}{s^2-1} - \coth^{-1} s - \frac{1}{3s^3(s^2-1)}\left(\frac{l}{ae}\right)^2 + \frac{2}{7s^7(s^2-1)}\left(\frac{l}{ae}\right)^6\right] \quad (24)$$

and substituting for s in terms of y

$$G = \frac{8\pi^2 b^2 I l n_1}{e}\left\{[1+(y/ae)^2]^{\frac{1}{2}} - (y/ae)^2 \operatorname{cosech}^{-1}(y/ae) - \frac{1}{3[1+(y/ae)^2]^{\frac{1}{2}}}\left(\frac{l}{ae}\right)^2 + \frac{2}{7[1+(y/ae)^2]^{\frac{7}{2}}}\left(\frac{l}{ae}\right)^6\right\} \quad (25)$$

The case of a thick coil of n_2 layers per unit depth of winding can now be considered by integration of this flux with respect to y. The resultant flux is then

$$K = n_2 \int_{r_1}^{r_2} G\,dy$$

where r_1 and r_2 are the inside and outside radii of the solenoid. Therefore

$$K = \frac{8\pi^2 b^2 I l n_1 n_2}{e}\int_{r_1}^{r_2}\left\{[1+(y/ae)^2]^{\frac{1}{2}} - (y/ae)^2 \operatorname{cosech}^{-1}(y/ae) - \frac{1}{3[1+(y/ae)^2]^{\frac{1}{2}}}\left(\frac{l}{ae}\right)^2 + \frac{2}{7[1+(y/ae)^2]^{\frac{7}{2}}}\left(\frac{l}{ae}\right)^6\right\}dy$$

$$= 8\pi^2 b^2 I l n_1 n_2 a\left[L + M(l/ae)^2 + N(l/ae)^6\right]_{y=r_1}^{y=r_2} \quad (26)$$

where

$$L = \frac{1}{3}\frac{y}{ae}[1+(y/ae)^2]^{\frac{1}{2}} + \frac{2}{3}\sinh^{-1}(y/ae) - \frac{1}{3}(y/ae)^3 \operatorname{cosech}^{-1}(y/ae) \quad (27)$$

$$M = \frac{-y/ae}{3[1+(y/ae)^2]^{\frac{1}{2}}} \quad (28)$$

$$N = \frac{2}{105}\left\{\frac{y/ae}{[1+(y/ae)^2]^{\frac{1}{2}}}\right\}\left\{\frac{3}{[1+(y/ae)^2]^2} + \frac{4}{1+(y/ae)^2} + 8\right\} \quad (29)$$

NUMERICAL VALUES

Equations (26–29) show that the flux through a coil of finite size may be calculated in terms of the three variables L, M and N which can be expressed in terms of the reduced radius y/ae, and are not dependent on the shape of the ellipsoid. The variables have been evaluated over a range of values of y/ae from 0 to 1·0, and the table shows the calculated values.

The dependence of L, M and N on y/ae

y/ae	0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
L	0	0·0990	0·1914	0·2843	0·3685	0·4470	0·5200	0·5880	0·6512	0·7101	0·7652
$-M$	0	0·0332	0·0654	0·0958	0·1238	0·1491	0·1715	0·1911	0·2082	0·2230	0·2357
N	0	0·0282	0·0546	0·0777	0·0968	0·1121	0·1231	0·1315	0·1375	0·1418	0·1448

Fig. 2 shows how these variables depend on y/ae. Initially they are all nearly proportional to y/ae, and this should be expected since the flux enclosed by turns near the surface of the ellipsoid will be nearly proportional to the number of turns. At higher values of y/ae, the variables all fall away from proportionality, and M tends to a minimum of $-1/3$, while N tends to 16/105. L does not tend to a maximum value, consequently for larger values of y/ae the latter terms in the series become of less importance.

APPLICATIONS

Using the expressions derived above, it should be possible to measure the changes of intensity of magnetization of an ellipsoid of revolution magnetized uniformly in the direction of its axis. For large ellipsoids it is probable that only a few layers of winding will be needed to give sufficient experimental accuracy, in which case a higher arithmetical accuracy will be obtained by using equation (25), or equation (13) if the coil extends over more than half the specimen. For ellipsoids of small dimensions the depth of winding will be much larger, and equations (26–29) must then be used.

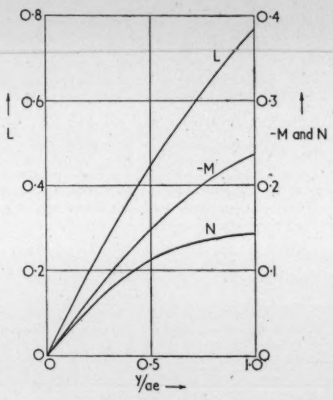


Fig. 2. The dependence of L, M and N on y/ae

ACKNOWLEDGEMENTS

The author is indebted to the Council of the Durham Colleges for the provision of a research studentship. He also wishes to express his thanks to Dr. W. D. Corner for his constant guidance and encouragement.

REFERENCES

(1) PEAKE, H. J., and DAVY, N. *Brit. J. Appl. Phys.*, **4**, p. 207 (1953).
(2) LIVENS, G. H. *The Theory of Electricity*, p. 256 (London: Cambridge University Press, 1918).
(3) STRATTON, J. A. *Electromagnetic Theory*, p. 211 (New York: McGraw Hill Book Co. Inc., 1941).
(4) PRACHE, P. M., and CAZENAVE, R. *Câbles et Transm.*, **4**, p. 216 (1950).

Some consideration of the errors of brightness and two-colour types of spectral radiation pyrometer

By E. C. Pyatt, B.Sc.(Eng.), A.M.I.E.E., National Physical Laboratory, Teddington, Middlesex

[*Paper received 5 October*, 1953]

Expressions are derived for the errors of the two types and from them a chart, from which these errors can be calculated for any pyrometer for all possible values of emissivity. The limitations of both types for direct measurement of true temperature are discussed and the errors likely to result when true temperatures are calculated from apparent temperature from previous knowledge of emissivities.

An expression is derived for the relation between the emissivities (ϵ_1 and ϵ_2) at the two-colour pyrometer wavelengths (λ_1 and λ_2) corresponding to a given ratio Z between the error of the two-colour pyrometer and the error of a brightness pyrometer working at one of the wavelengths (λ_2). This information is plotted as a chart of ϵ_1 against ϵ_2, the application of which to a practical problem is demonstrated.

A more complex type of ratio pyrometer is put forward, enabling broader assumptions to be made regarding the emissivity/wavelength relation.

1. General

The brightness type of spectral radiation pyrometer measures the amount of energy being radiated by a hot body in a limited wavelength range, which by Wien's law is a measure of the temperature of the body. The measurement is made either by comparing the brightness of the hot body with that of a variable source in the instrument (disappearing filament optical pyrometers; certain comparison photoelectric methods, etc.) or directly with a photoelectric cell. It is well known that this type of pyrometer only measures true temperature when sighted on a black-body. At any particular wavelength the energy radiated by a non-black-body is always less than that radiated by a black-body at the same temperature; the ratio between these two energy values is known as the spectral emissivity at that wavelength and is always less than unity. In the case of non-black-bodies, therefore, the brightness temperature pyrometer will read an apparent temperature less than the true temperature and the value of the emissivity must be known before the value of the latter can be deduced from the observations.

To overcome this objection the two-colour pyrometer was put forward. This measures the ratio between the amounts of energy radiated by the hot body in two limited wavelength ranges. If the emissivity of the body is the same at these two wavelengths, a so-called grey-body, then the pyrometer will read the true temperature of the body regardless of the actual value of the emissivity, which cancels out when the ratio is taken. For non-grey-bodies the two-colour pyrometer will read an apparent temperature which may be greater or less than the true temperature, depending on the values of the emissivities at the two wavelengths, which must be known, therefore, before the value of the true temperature can be deduced from the observations.

In the following work it is proposed to use, instead of "two-colour pyrometer," the term "ratio pyrometer," which more exactly describes its function. In practice, wide variations occur in the values of emissivity and slope of the emissivity/wavelength characteristic and little is known of the respective merits of the two types of pyrometer. One of the purposes of this paper is to compare the behaviour of the two types under all possible emissivity conditions and to provide a simple means for estimating the error of either in any particular case. It is assumed that the pyrometers are sensitive to a single wavelength, so that the energy measured is given by the expression for Wien's law at that wavelength. This approximates closely to the practical case where only a narrow band of wavelengths is used.

2. Derivation of the Errors

We consider a hot-body having a true temperature $T°$ K with emissivities ϵ_B, ϵ_1 and ϵ_2 at wavelengths λ_B, λ_1 and λ_2 respectively. As Wien's law involves the reciprocal of the temperature it is more convenient first of all to discuss the error in $1/T$, which is converted to an error in T at a later stage.

If S_B is the apparent temperature of the hot-body as read by a brightness pyrometer working at a wavelength of λ_B then

$$\epsilon_B C_1 \lambda_B^{-5} \exp(-C_2/\lambda_B T) = C_1 \lambda_B^{-5} \exp(-C_2/\lambda_B S_B) \quad (1)$$

$$\log_e \epsilon_B = \frac{C_2}{\lambda_B}\left(\frac{1}{T} - \frac{1}{S_B}\right) \quad (2)$$

Now writing $(1/T) - (1/S_B)$ as E_B (the error in $1/T$ for the brightness pyrometer)

$$E_B = (\lambda_B \log_e \epsilon_B / C_2) \quad (3)$$

or

$$E_B = K_B \log_e \epsilon_B \text{ (where } K_B = \lambda_B/C_2) \quad (4)$$

This is the brightness pyrometer error equation.

If S_R is the apparent temperature of the hot-body as read by a ratio pyrometer working at wavelengths of λ_1 and λ_2 then

$$\frac{\epsilon_1 C_1 \lambda_1^{-5} \exp\left(-\frac{C_2}{\lambda_1 T}\right)}{\epsilon_2 C_1 \lambda_2^{-5} \exp\left(-\frac{C_2}{\lambda_2 T}\right)} = \frac{C_1 \lambda_1^{-5} \exp\left(-\frac{C_2}{\lambda_1 S_R}\right)}{C_1 \lambda_2^{-5} \exp\left(-\frac{C_2}{\lambda_2 S_R}\right)} \quad (5)$$

$$-\log_e \epsilon_1/\epsilon_2 = C_2\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right)\left(\frac{1}{T} - \frac{1}{S_R}\right) \quad (6)$$

Now writing $(1/T) - (1/S_R)$ as E_R (the error in $1/T$ for the ratio pyrometer)

$$E_R = -\frac{\lambda_1 \lambda_2}{C_2(\lambda_1 - \lambda_2)} \log_e \epsilon_1/\epsilon_2 \quad (7)$$

or

$$E_R = K_R K_1 \log_e \epsilon_1/\epsilon_2 \quad (8)$$

where $K_1 = \lambda_1/(\lambda_2 - \lambda_1)$ and $K_R = \lambda_2/C_2$.

This is the ratio pyrometer error equation.

3. GRAPHICAL SOLUTION FOR THE ERRORS

A simple nomogram of three parallel straight lines (scales 1, 2 and 3 on Fig. 1) linear respectively on $\log\epsilon_2$, $\log\epsilon_1$ and $\log\epsilon_1/\epsilon_2$ is used to obtain the value of the latter function. The three lines are equally spaced and the scales of $\log\epsilon_2$ and $\log\epsilon_1$ are designated in ϵ_2 and ϵ_1. A straight line connecting two values of ϵ_2 and ϵ_1 on Fig. 1 gives the value of $\log\epsilon_1/\epsilon_2$ on scale 3.

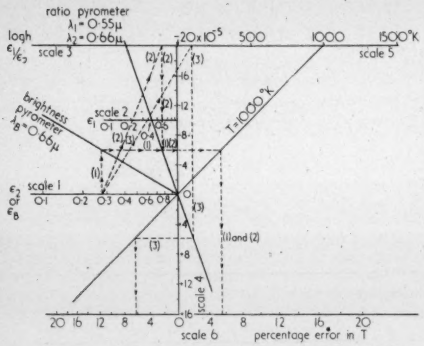


Fig. 1. The complete nomogram (with examples)

$\lambda_1 = 0{\cdot}55\,\mu$; $\lambda_2 = 0{\cdot}66\,\mu$; $T = 1000°$ K.

(1) brightness pyrometer, $\epsilon_B = 0{\cdot}28$; (2) ratio pyrometer, $\epsilon_1 = 0{\cdot}21$, $\epsilon_2 = 0{\cdot}28$; (3) ratio pyrometer, $\epsilon_1 = 0{\cdot}35$, $\epsilon_2 = 0{\cdot}28$

The error equations derived above are of similar form. The graphs of E_B against $\log\epsilon_B$ and of E_R against $\log\epsilon_1/\epsilon_2$ are straight lines through the origin having slopes of K_B and K_RK_1 respectively. Scales 1 and 3 are parallel scales of $\log\epsilon_B$ and $\log\epsilon_1/\epsilon_2$ in the same units and having the same zero; it is therefore possible to solve both error equations by adding one error scale (scale 4) and drawing through the zero of the scale a straight line of slope K_B for the brightness pyrometer error equation or a straight line of slope K_RK_1 for the ratio pyrometer error equation. The brightness pyrometer line ends at zero as the error cannot be positive; the ratio pyrometer line extends on either side of zero as the error is positive or negative depending on the slope of the emissivity/wavelength characteristic.

The method of deriving the errors of a brightness and a ratio type pyrometer having a common wavelength from the full nomogram is shown in Fig. 1 for one value of ϵ_2 in combination with two different values of ϵ_1. The values of ϵ_1 and ϵ_2 are connected by a straight line, which is produced to cut the $\log\epsilon_1/\epsilon_2$ scale. A perpendicular from ϵ_2 to the brightness pyrometer line gives the brightness pyrometer error. A perpendicular from the derived value on the $\log\epsilon_1/\epsilon_2$ scale on to the ϵ_2 scale gives the ratio pyrometer error as an intercept between the ratio pyrometer line and the ϵ_2 scale.

The method is, of course, applicable to all brightness and ratio type pyrometers whatever the working wavelengths.

4. PERCENTAGE ERRORS IN TEMPERATURE

The errors of both types of pyrometer were defined by similar expressions, which may be written in general:

$$(1/T) - (1/S) = E \tag{9}$$

We now define the corresponding error in temperature by

$$S - T = T\sigma \tag{10}$$

(where 100σ is the percentage error in T).

Eliminating S gives

$$E = \sigma/[T(1+\sigma)] \tag{11}$$

It can be seen that the graph of E against $\sigma/(1+\sigma)$ is a straight line through the origin having a slope of $1/T$.

In Fig. 1, scale 5, which is linear in T and parallel to the $\log\epsilon_2$ scale, enables lines of slope $1/T$ to be drawn through the origin. A "percentage error in T" scale (scale 6) is also drawn, nominally through the origin, but actually lower down the figure to avoid confusion. This scale is divided in $\sigma/(1+\sigma)$ but is designated directly in σ. The extra construction lines enabling the percentage error in T to be determined at a given temperature for the examples of Section 3 have also been added in Fig. 1.

5. LIMITATIONS ON DIRECT MEASUREMENT

The range of values of emissivity for which the percentage error for a given type of pyrometer will not exceed a given value can be easily shown.

For the brightness pyrometer (from Section 4)

$$E = K_B \log\epsilon_B = \sigma_1/[T(1+\sigma_1)] \tag{12}$$

(where $100\sigma_1$ is the maximum permitted value of the percentage error—assumed to be small).

Whence

$$\log\epsilon_B = -\sigma_1/K_BT \tag{13}$$

(σ is negative for the brightness pyrometer).

Therefore for a brightness temperature pyrometer to have a percentage error less than $100\sigma_1$ the condition

$$0 > \log\epsilon_B > -\sigma_1/K_BT \tag{14}$$

must be fulfilled.

For the ratio pyrometer similarly

$$E = K_RK_1 \log\epsilon_1/\epsilon_2 = \sigma_1/[T(1+\sigma_1)] \tag{15}$$

Whence

$$\log\epsilon_1/\epsilon_2 = \sigma_1/K_RK_1T \tag{16}$$

Therefore for a ratio pyrometer to have a percentage error less than $100\sigma_1$ the condition

$$\log\epsilon_2 + \frac{\sigma_1}{K_RK_1T} > \log\epsilon_1 > \log\epsilon_2 - \frac{\sigma_1}{K_RK_1T} \tag{17}$$

i.e.
$$\epsilon_2\left(1+\frac{\sigma_1}{K_RK_1T\log\epsilon_2}\right) > \epsilon_1 > \epsilon_2\left(1-\frac{\sigma_1}{K_RK_1T\log\epsilon_2}\right) \tag{18}$$

must be fulfilled.

6. TRUE TEMPERATURE BY CALCULATION

When the values of emissivity are such that the observed temperature is different from the true temperature by more than a given percentage, it may be necessary to calculate the

latter from the former by the formulae of Section 2. In assessing the relative merits of the two pyrometer types in this instance, we have to investigate the effect of errors in our knowledge of the emissivity values and the relative sensitivities.

We can express an uncertainty in the value of emissivity ϵ by the expression

$$\epsilon' = \epsilon(1 + \alpha) \tag{19}$$

where ϵ' is the value of emissivity believed to be correct.

Let T be the true temperature and T' the value of the true temperature calculated on the value ϵ' for the emissivity.

For the brightness pyrometer

$$\frac{1}{T} - \frac{1}{S_B} = K_B \log_{10} \epsilon_B \tag{20}$$

and

$$\frac{1}{T'} - \frac{1}{S_B} = K_B \log_{10} \epsilon_B(1 + \alpha) \tag{21}$$

whence

$$T - T' = TT'K_B \log_{10} (1 + \alpha) \tag{22}$$

Now write

$$T' = T(1 - \beta') \tag{23}$$

where $100\beta'$ is the percentage error in T.

$$T\beta' = T^2(1 + \beta')K_B \log_{10} (1 + \alpha) \tag{24}$$

or for β' and α both small

$$\beta' = TK_B\alpha \tag{25}$$

(This is negligible in most practical cases, i.e. the error is $\frac{1}{2}°$ at $T = 1000°$ K, $\lambda = 0{\cdot}66\ \mu$, $\alpha = 1/100$.)

For the ratio pyrometer similarly if T'' is the value of temperature (corresponding to a percentage error in T of β'') calculated from values of $\epsilon_1' = \epsilon_1(1 \pm \alpha)$ and $\epsilon_2' = \epsilon_2(1 \pm \alpha)$, then for β'' and α both small it can be shown that β'' lies between $\pm TK_RK_1 2\alpha$.

If the percentage uncertainties in ϵ_1 and ϵ_2 are of the same magnitude and sign then β'' will be zero. If of the same magnitude and opposite sign, the error in the value of the true temperature calculated from the observed ratio temperature will be greater than that calculated from the observed brightness temperature by a factor of $2K_RK_1/K_B$ (or $2K_1$ when the pyrometers have a common wavelength).

(For a brightness pyrometer working at $0{\cdot}66\ \mu$ and a ratio pyrometer working at $0{\cdot}55$ and $0{\cdot}66\ \mu$ the factor would be 10 and the error introduced into the calculation would not be negligible.)

The relative sensitivity of the two types can be determined by differentiating the error equations in Section 2.

For the brightness pyrometer

$$-dT/T^2 = -dS_B/S_B^2 \quad \text{or} \quad dS_B/dT = S_B^2/T^2 \tag{26}$$

for the ratio pyrometer similarly

$$dS_R/dT = S_R^2/T^2 \tag{27}$$

so that

$$\frac{dS_B}{dS_R} = \frac{S_B^2}{S_R^2} \tag{28}$$

This slightly favours the ratio pyrometer. On the other hand the accuracy of determination of the apparent temperature is likely to be inferior in the case of the ratio pyrometer, which involves the more complicated process of measuring two brightnesses and then presenting the ratio between them.

7. THE RATIO OF THE ERRORS

In this section it is assumed that the pyrometers have a common wavelength, i.e. that in Section 2:

$$K_B = K_R(= K) \quad \text{or} \quad \lambda_B = \lambda_2 \tag{29}$$

For the brightness pyrometer we define the error in temperature (D_B) by the expression

$$D_B = \sigma_B T = S_B - T \tag{30}$$

For the ratio pyrometer similarly

$$D_R = \sigma_R T = S_R - T \tag{31}$$

and we write

$$D_R/D_B = Z \tag{32}$$

i.e. Z is the ratio of the errors in temperature.

From Section 2 we quote the error equations

$$\frac{1}{T} - \frac{1}{S_B} = K \log_{10} \epsilon_2$$

and

$$\frac{1}{T} - \frac{1}{S_R} = KK_1 \log_{10} \epsilon_1/\epsilon_2$$

Eliminating S_B and S_R we obtain a relation between $\log_{10} \epsilon_1$ and $\log_{10} \epsilon_2$ involving Z the ratio between the errors of the two pyrometer types. This is

$$\log_{10} \epsilon_1 = \log_{10} \epsilon_2 \left\{ \frac{1}{K_1} \left[TK \log_{10} \epsilon_2 + \frac{1}{Z}(1 - TK \log_{10} \epsilon_2) \right]^{-1} + 1 \right\} \tag{33}$$

Z is positive or negative depending on whether D_R is negative or positive. The range of values of emissivity for which the ratio pyrometer error is less than the brightness pyrometer error by a factor Z is given by

$$\log_{10} \epsilon_2 \left\{ \frac{1}{K_1} \left[TK \log_{10} \epsilon_2 + \frac{1}{Z}(1 - TK \log_{10} \epsilon_2) \right]^{-1} + 1 \right\} < \log_{10} \epsilon_1 < \log_{10} \epsilon_2 \left\{ \frac{1}{K_1} \left[TK \log_{10} \epsilon_2 - \frac{1}{Z}(1 - TK \log_{10} \epsilon_2 \right]^{-1} + 1 \right\} \tag{34}$$

For equal errors, i.e. $Z = 1$ this reduces to

$$\log_{10} \epsilon_2 \left(\frac{1}{K_1} + 1 \right) < \log_{10} \epsilon_1 < \log_{10} \epsilon_2 \left[\frac{1}{K_1} \left(\frac{1}{2TK \log_{10} \epsilon_2 - 1} \right) + 1 \right] \tag{35}$$

or

$$\epsilon_2^{\left(1 + \frac{1}{K_1}\right)} < \epsilon_1 < \epsilon_2^{\left[\frac{1}{K_1}\left(\frac{1}{2TK \log_{10} \epsilon_2 - 1}\right) + 1\right]} \tag{36}$$

8. APPLICATION TO PRACTICAL PROBLEMS

The nomogram of Sections 3 and 4 is easily constructed and its value for making calculations involving errors of either type of pyrometer working at any wavelength is obvious.

The information of Section 5 is inherent in the nomogram, but sometimes it may be simpler to draw a small scale of T for a given value of σ, or a small scale of σ for a given value of T, coincident with the $\log_{10} \epsilon_1/\epsilon_2$ scale (scale 3) or the $\log_{10} \epsilon_2$

scale (scale 1) so that the limits for a particular pyrometer can be more rapidly assessed.

Section 5 also enables a chart of ϵ_1 against ϵ_2 to be drawn for a particular ratio pyrometer, using an appropriate maximum value of T for the problem in hand. For one particular value of σ, this is of the form shown in Fig. 2.

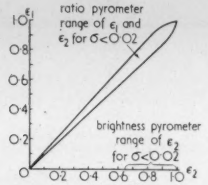


Fig. 2. Relation between ϵ_1 and ϵ_2 for $\sigma < 0{\cdot}02$ at:
$T = 1000°$ K; $\lambda_1 = 0{\cdot}55\,\mu$; $\lambda_2 = 0{\cdot}66\,\mu$.

The condition for any hot-body (emissivity ϵ_1 at λ_1 and ϵ_2 at λ_2) can be located as a point on this chart and if the point is enclosed by the curves the ratio pyrometer error will be less than σ. Curves for other values of σ could readily be added, bearing in mind that σ must be small in order not to invalidate the assumptions of Section 5. The corresponding brightness pyrometer range is a short portion of the ϵ_2 scale as shown.

Section 7 enables a comparison to be made between a particular brightness pyrometer and a particular ratio pyrometer having a common wavelength. The result can again be expressed in the form of a chart of ϵ_1 against ϵ_2. The general principle is shown in Fig. 3(*a*) which has been drawn

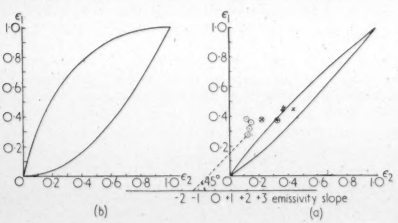


Fig. 3. Relation between ϵ_1 and ϵ_2 for $Z = 1$

(*a*) $\lambda_1 = 0{\cdot}55\,\mu$; $\lambda_2 = 0{\cdot}66\,\mu$.

○	copper	Burgess and Waltenburg*
+	nickel	
⊕	palladium	
⊗	gold	
×	tungsten	Hamaker*

(*b*) $\lambda_1 = 0{\cdot}55\,\mu$; $\lambda_2 = 1{\cdot}10\,\mu$.

for $Z = 1$, $\lambda_1 = 0{\cdot}55\,\mu$, $\lambda_2 = 0{\cdot}66\,\mu$ and $T = 3000°$ K. Again the conditions for any hot-body being measured can be located as a point on this chart. When the point lies inside the curves the ratio pyrometer has the lesser error, outside the curves the brightness pyrometer has the lesser error. Some values of emissivity published by various investigators have been added in Fig. 3(*a*). In Fig. 3(*b*) the same calculations have been made for wavelengths of $0{\cdot}55\,\mu$ and $1{\cdot}10\,\mu$; the area between the curves is considerably increased; and although a bigger difference between ϵ_1 and ϵ_2 would be likely it will not necessarily be proportionate. For lower temperatures the upper curve would be raised still further and the area further increased.

Fig. 4 shows a similar chart worked out for a limited range of temperatures and emissivities as met with in measurements

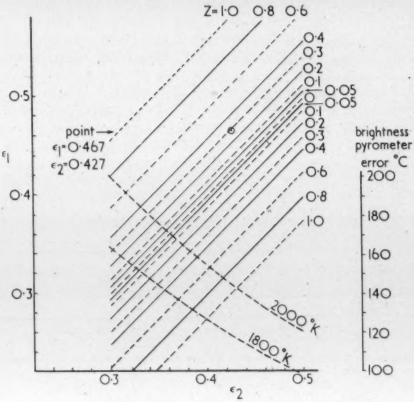


Fig. 4. Relation between ϵ_1 and ϵ_2 for $T = 2000°$ K
$Z = 0$ to 1; $\lambda_1 = 0{\cdot}467\,\mu$; $\lambda_2 = 0{\cdot}66\,\mu$.

on liquid steel. In this case $\lambda_1 = 0{\cdot}467\,\mu$. $\lambda_2 = 0{\cdot}66\,\mu$, $T = 2000°$ K the range of ϵ_2 is 0·3 to 0·5 and the range of Z is 0 to 1. A series of curves of the error of the brightness pyrometer against the value of ϵ_2 has been added so that the actual magnitude of the errors as well as the ratio between them can also be determined. Because a maximum value of true temperature has to be assumed before such a chart can be drawn, there is some loss of accuracy when the chart is used for a range of temperature.

A typical measurement of brightness temperature and ratio temperature on a plain carbon steel gave values of brightness temperature ($\lambda = 0{\cdot}66\,\mu$) and ratio temperature ($\lambda_1 = 0{\cdot}467\,\mu$ and $\lambda_2 = 0{\cdot}66\,\mu$) of 1395 and 1560° C respectively. A thermocouple gave the true temperature as 1520° C.* These figures lead to values of 0·427 and 0·467 for the emissivities at $\lambda = 0{\cdot}66\,\mu$ and $0{\cdot}467\,\mu$ respectively.

Plotting these particular emissivity values on Fig. 4, i.e. now assuming that the emissivities are known beforehand and the errors have to be calculated, we find that an error ratio of 0·29 is forecast. If the approximate value of the true temperature is known the brightness temperature pyrometer error can also be read off from the chart. This is 118° and the ratio temperature error is therefore forecast as 34°.

* *Temperature. Its Measurement and Control in Science and Industry*, pp. 1184, 1185 (New York: American Institute of Physics, 1941).

* I am indebted to Mr. J. A. Hall of the National Physical Laboratory for these observations which were obtained in the course of some work which has not yet been published.

These figures show reasonable agreement with the above observations—brightness pyrometer error 125°, ratio pyrometer error 40°, error ratio 0·32.

It should still be borne in mind, however, that where the magnitude of the error is such that the true temperature has to be calculated from the apparent temperature, the accuracy of the determination of the true temperature depends entirely on the accuracy, with which the emissivities are known, and on the accuracy of measurement of the apparent temperature, *not* on the magnitude of the error.

9. MORE COMPLEX TYPES OF RATIO PYROMETER

The above considerations suggest that measurements at further wavelengths might enable broader, and perhaps more valid, assumptions to be made regarding the emissivity/wavelength relation, in place of the equality of emissivity at two wavelengths demanded by the simple ratio pyrometer. It is shown below that if two ratio measurements are made between the amounts of energy radiated at three wavelengths, then the true temperature can be deduced from the

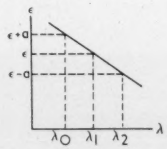


Fig. 5. Assumed relation between ϵ and λ

two observed apparent temperatures, provided that the emissivity/wavelength relation is linear for these three wavelengths and the sign of its slope is known. It is not necessary to know the emissivity values. Fig. 5 shows the relation between the quantities involved.

For a ratio pyrometer working at λ_0 and λ_1 we can write from Section 1

$$\frac{1}{T} - \frac{1}{S_{\lambda 01}} = K_{01} \operatorname{logh}\left(\frac{\epsilon + a}{\epsilon}\right) \tag{37}$$

$$\left(\text{where } K_{01} = \frac{\lambda_1}{C_2} \cdot \frac{\lambda_0}{\lambda_1 - \lambda_0}\right).$$

Similarly for a pyrometer working at λ_1 and λ_2

$$\frac{1}{T} - \frac{1}{S_{\lambda 12}} = K_{12} \operatorname{logh}\left(\frac{\epsilon}{\epsilon - a}\right) = -K_{12} \operatorname{logh}\left(\frac{\epsilon - a}{\epsilon}\right) \tag{38}$$

$$\left(\text{where } K_{12} = \frac{\lambda_2}{C_2} \cdot \frac{\lambda_1}{\lambda_2 - \lambda_1}\right).$$

Eliminating $1/T$

$$K_{01} \operatorname{logh}\left(1 + \frac{a}{\epsilon}\right) + K_{12} \operatorname{logh}\left(1 - \frac{a}{\epsilon}\right) = \frac{1}{S_{\lambda 12}} - \frac{1}{S_{\lambda 01}} = D_\lambda \tag{39}$$

The true temperature is derived from the observed apparent temperatures as follows:

For a particular double ratio pyrometer, i.e. for given values of λ_0, λ_1 and λ_2, a curve is plotted of $K_{01} \operatorname{logh}(1 + a/\epsilon) + K_{12} \operatorname{logh}(1 - a/\epsilon)$ against $K_{01} \operatorname{logh}(1 + a/\epsilon)$ for all possible values of a/ϵ (i.e. $0 < a/\epsilon < 1$). We are then assuming that the slope of the emissivity/wavelength characteristic is such that emissivity decreases as wavelength increases. For the reverse case, which is not so frequently met, the plot would need to be continued from 0 to -1 and there would then be two values of the function corresponding to any observed value of D_λ. The sign of the slope must therefore be known if we are to avoid ambiguity.

From this curve the value of $K_{01} \operatorname{logh}(1 + a/\epsilon)$ can be derived from the value of D_λ obtained from the observations. The value of T is now deduced from the equation $K_{01} \operatorname{logh}(1 + a/\epsilon) = (1/T) - (1/S_{\lambda 01})$. This last step is readily carried out by a nomogram of three straight lines and the whole solution reduces to the simple process illustrated in Fig. 6.

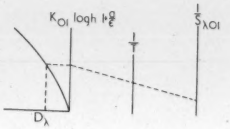


Fig. 6. Graphical solution of double ratio pyrometer

The proposed double ratio pyrometer appears to offer a considerable improvement on the single ratio type for the direct measurement of the true temperature. A discussion of alternative methods and an estimation of the errors of this type of pyrometer will form the subject of a future paper.

ACKNOWLEDGEMENTS

The work described above has been carried out as part of the research programme of the National Physical Laboratory and the paper is published by permission of the Director of the Laboratory.

Correspondence

The molecular flow of gases

The paper by Rossmann and Yarwood published in the January 1954 issue of this *Journal* provides an excellent example of the analogous aspects of energy radiation and molecular flow phenomena. Geometric or configuration factors have been calculated for numerous systems of radiators and receivers, in both the fields of heat transfer and illumination engineering. These may well be used as the starting points of solutions to many molecular flow problems, and where computation is lengthy, established model experimental techniques may be used.*

With the use of the two theorems given in the forementioned paper, it is possible to derive the celebrated "long tube" molecular streaming formula of Knudsen.

Consider two equi-circular parallel coaxial apertures of radius r separated by a distance z, where these lengths are small compared with the mean free path of the gas molecules which leave a large constant temperature enclosure *via* one of these apertures with velocity components towards the other. The fraction of these molecules which cross the outer aperture is given by

$$1 + (z^2/2r^2)[1 - (1 + 4r^2/z^2)^{\frac{1}{2}}] = f(z/r)$$

If now the inner aperture forms the entrance to a cylindrical tube of the same radius, the fraction of those molecules entering the tube which make a first collision with the wall of the tube between distances z and $z + dz$ from the entrance is $-f'(z/r) \,.\, dz$ and so the average distance from the entrance at which the first collision is made will be $-\int_0^\infty f'(z/r) \,.\, z \,.\, dz$, where the upper limit of the integral has been chosen as ∞ since the tube length l is taken to be very large compared with r.

The above integral very simply yields the value $4r/3$. It is readily seen that molecules will make successive collisions with the tube walls at a mean distance of separation of twice the value found by reference to the entrance aperture for the first collision.

If the tube is long, then the molecules may be considered to make their first collision at a distance of $4r/3$ from the entrance aperture, and for those that proceed down the tube, the next collision will be at a further distance $8r/3$ down the tube and so on; the last collision being at a distance of $4r/3$ from the exit of the tube. There will therefore be $N = 3l/8r$ sections in the tube at which collisions between gas molecules and wall will occur. If at each collision the molecules have equal chances of moving off the wall towards entrance or exit, it is easily shown that the fraction $1/(1 + N)$ of those which originally enter the tube, emerge from the exit. By substitution, this fraction becomes $8r/(8r + 3l)$, which applied to the total number rate at which molecules cross the entrance aperture into the tube, yields the Knudsen expression.

When the tube is "short", the averaging process adopted breaks down. For example, a tube of length equal to half its radius allows about 61% of the molecules to escape without collision with its walls, and then the average distance from the entrance at which the other 39% of the molecules make a first collision with the wall is just under $r/4$. The average distance at which molecules strike the wall for the second and succeeding times will differ. In the case cited where those that do collide with the tube strike it at an average distance half way along its length, it might be inferred that half of these finally escape from the exit and half return across the entrance. On this basis roughly 80% of those molecules incident on the entrance aperture go through the tube. Thus the numeric in the Knudsen expression is amended from 3/8 to about 1/2.

There is one small criticism of the Rossmann and Yarwood paper, which pertains to the remarks concerning their equation (17). Modifications to the Knudsen or Clausing expressions for condensation coefficients of less than unity are easily deduced, as indeed shown by their analysis from expression (15) onwards.

Department of Applied Physics,
Northampton Polytechnic,
London, E.C.1.

D. W. STOPS

* ECKERT, E. R. G. *Introduction to the Transfer of Heat and Mass*, p. 211 (New York: McGraw-Hill Book Co. Inc., 1950).

A check on the standard observer data at 4358 Å

Since the publication of my paper in the March 1953 issue of this *Journal*, Dr. Judd of the National Bureau of Standards has drawn my attention to a more recent proposal for modification of the standard observer data than the Wald–Gibson–Tyndall modification used in my paper. The proposal is that given in Table 2 of the Secretariat Report No. 7 of the 12th C.I.E. meeting, Stockholm, 1951, and incorporates the Wald–Gibson–Tyndall and other data adjusted to satisfy essential requirements of the C.I.E. system with regard to spectrum colours and standard illuminant *B*.

Whereas the Wald–Gibson–Tyndall data as used in my paper gave an unsatisfactory correlation between visual measurements and computed values of chromaticity, the $X_1\ Y_1\ Z_1$ data of the newer proposal gives slightly better correlation than does the C.I.E. 1931 data.

The various chromaticities and factors obtained with the $X_1\ Y_1\ Z_1$ data are given below.

Table 1. *Chromaticities of standard and other illuminants using* $X_1\ Y_1\ Z_1$ data

	x_1	y_1	z_1
Illuminant *A*	0·449	0·412	0·139
Illuminant *B*	0·351	0·357	0·292
Illuminant *C*	0·313	0·321	0·366
2500° K radiator	0·473	0·422	0·105
Green fluorescent	0·234	0·604	0·162

Table 2. *Multiplying factors to convert 8-band luminance from C.I.E. to* Y_1

Band No.	*Warm-white fluorescent*	*Daylight fluorescent*
1	7·7	6·9
2	1·9	1·95
3	1·18	1·19
4 to 8	1·0	1·0

Table 3. *Factors for obtaining* $\Sigma E_\lambda \bar{x}_{1\lambda}$ *and* $\Sigma E_\lambda \bar{z}_{1\lambda}$ *from* Y_1 *luminance data*

(The luminance is multiplied by the factor shown)

Band No.	Warm-white		Daylight	
	$\Sigma E_\lambda \bar{x}_{1\lambda}$	$\Sigma E_\lambda \bar{z}_{1\lambda}$	$\Sigma E_\lambda \bar{x}_{1\lambda}$	$\Sigma E_\lambda \bar{z}_{1\lambda}$
1	13·5	63	13·5	63
2	9·9	47	10·1	50
3	5·8	29·7	5·85	30·0
4	0·39	3·14	0·41	3·1
5	0·38	0·03	0·34	0·03
6	1·16	0·002	1·19	0·002
7	2·24		2·24	
8	2·7		2·7	

Table 4. *Chromaticities obtained for the warm-white and daylight fluorescent lamps of original paper when above data is applied to visual and spectral band measurements*

	x_1	y_1	z_1
Warm-white from spectral bands	0·423	0·394	0·183
Warm-white from visual	0·422	0·390	0·188
difference spectral bands from visual	+0·001	+0·004	−0·005
Daylight from spectral bands	0·366	0·363	0·271
Daylight from visual	0·362	0·366	0·272
difference spectral bands from visual	+0·004	−0·003	−0·001

Lamp Research Laboratory, W. HARRISON
Siemens Electric Lamps and Supplies Ltd.,
Preston.

Line interaction in X-ray and electron diffraction photographs

Mr. Archard's contribution published in the January issue of this *Journal* (p. 19) omits one subjective factor which may be of importance. If we observe a complex line having the profile of Fig. 1, then we *seem* to see two lines in the positions indicated by *A* and *B*, although in fact only one peak is present in the blackening curve. The illusory peak is situated towards the steeper fall of the density.

It may also be observed that if a single unsymmetrical peak is present as in Fig. 2, the eye will see the peak displaced towards the steeper side (in some such position as *A*), so that different peak positions will be obtained by direct observation, and by measurement of a microphotometer record. This type of illusion has been discussed by Aruja.*

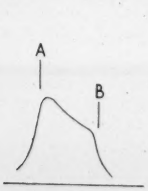


Fig. 1. "Subjective" repulsion of overlapping peaks

Fig. 2. "Subjective" displacement of asymmetrical peak

If Mr. Archard's measurements were made directly on his photographs and not on a microphotometer record, then this illusion should also be considered in interpreting his Fig. 4. If two peaks are overlapping to a considerable extent, so as to produce a resultant similar to my Fig. 1, then an apparent repulsion of the peaks will occur. This repulsion will diminish when the peaks separate, leaving only the genuine effect of attraction. Of course the developer effect mentioned by Mr. Archard may be present as well.

Pedology Department, DOUGLAS M. C. MACEWAN
Rothamsted Experimental Station,
Harpenden, Herts.

I am most grateful to Dr. MacEwan for pointing out a subjective contributory cause for the observations reported, and for drawing attention to the work of Aruja, which had escaped my notice. The X-ray photographs were in fact measured directly; the electron diffraction photographs, which exhibited only the attractive effect, were measured by means of a scanning microphotometer, which would not be subject to the illusion mentioned.

The University, G. D. ARCHARD†
Reading.

* ARUJA. *Nature* [*London*], **154**, p. 53 (1944).

† Now at: Associated Electrical Industries Ltd., Aldermaston, Berks.

The firing of pyrophyllite for minimum distortion

At some temperature during the firing of the naturally occurring clay mineral known as pyrophyllite, a change of crystalline structure results, and severe distortion may develop. Since most of that distortion takes place at low firing temperatures, a pre-baking method has been found generally useful in avoiding excessive and irregular contractions or expansions in different batches of the material. Prior to machining, the pyrophyllite is fired embedded in French chalk (in a vertical position if it is a lamina) by heating it slowly over a period of two hours to 750° C, keeping at that temperature overnight and then allowing it to cool in the furnace. The semi-baked material is still sufficiently soft to be machined and then, to produce the normal hard surface, it may again be fired as before, but at 1050° C. The average change in dimensions (increase or decrease) occurring when this technique is used is 0·08 ± 0·08%; about twenty times smaller than when simple firing is employed.

Research Laboratory, V. E. DYKE
Associated Electrical Industries Ltd.,
Aldermaston, Berks.

The noise of valves

In my paper published on page 127 of the April issue of this *Journal*, I allowed to pass in proof reading a series of errors in the section entitled "Estimates of $\overline{(\partial i/\partial V)^2}$ and $\overline{(\partial i/\partial i_s)^2}$."

In this section α should replace ϵ in the following places: lines 6, 7, 9, 19 (l.h.s. of equation), 25, 28, 29 (l.h.s. of equation), 37 (l.h.s. of equation).

The error was caused by choosing the same letter from two different founts of Greek for related entities.

Electronic Tubes Ltd., C. S. BULL
High Wycombe, Bucks.

New books

Monographies C, Physique Mathematique: 11. Les fonctions orthogonales dans les problems aux limites de la physique mathematique. By T. VOGEL. (Paris: Centre National de la Recherche Scientifique.) Pp. 191 + iv. Price 30*s*. 7*d*.

This monograph is written in French for physicists and engineers who are concerned with the differential equations of mathematical physics. Such equations are frequently solved by the separation of variables and the subsequent development of solutions in series of orthogonal functions. The general theory of orthogonal functions, resting on a rigorous discussion of orthogonality and completeness, is given in Chapter 1. The following chapter deals with the more important orthogonal functions (Fourier, Bessel, Legendre, Mathieu and others) and the last with examples of the use of these functions in physical problems.

The attack on this subject is both direct and elegant. Very extensive references are given to recent developments, and in all this book is instructive and pleasant to read.

I. I. BERENBLUT

Flying saucers. By DONALD H. MENZEL. (Cambridge, Mass., U.S.A.: Harvard University Press; London: Putnam and Co. Ltd.) Pp. xii + 319. Price 21*s*.

In this book Professor Menzel examines critically the evidence relating to the occurrence and nature of the various apparitions, especially prevalent in U.S.A. and commonly designated "flying saucers." Admittedly the evidence usually leaves much to be desired, even in the relatively few cases when the observations were made by professional scientists. Moreover, the whole subject is bedevilled by public credulity and love of sensation, by press exploitation and, in some cases, by deliberate invention. Nevertheless, it is clear that there are a great many observations which deserve serious investigation, and these are discussed very thoroughly by the author. In doing so he ranges over a wide field, mainly in atmospheric optics, but including also the aurora, terrestrial magnetism, radar, comets and other astronomical phenomena. Naturally it is not possible under the circumstances to account with complete certainty for even a fraction of the apparitions reported, but one can at least say that Dr. Menzel's suggested explanations are highly plausible, and that he fully establishes his case that there is no reason to suppose that any animate extra-terrestrial agency is involved. It is an unusual, entertaining and instructive book, of special interest to the physicist, partly as a model of dispassionate investigation in a difficult field, but also as a review of many of the less familiar optical effects associated with our atmosphere. W. E. CURTIS

The theory of metals (2nd edition). By A. H. WILSON, M.A., F.R.S. (London: Cambridge University Press.) Pp. vi + 346. Price 45*s*.

"The first edition," it is stated on the jacket of this book, "appeared in 1936, and has held its place ever since as the best large-scale treatment of the subject". Whatever reservations may be felt about this pronouncement, there is no doubt that the second edition, as a book, is very much better than the first. Although much longer, it forms a more coherent whole, partly through the omission of certain topics, such as surface effects, which involve physical ideas different from those forming the large main theme; greater consideration is given to the reader in the manner of presentation; over the field covered, a more balanced impression is given of the present state of the subject, and the references, at the end of the chapters, are a carefully considered selection of key papers and review articles rather than a random sample; and, in contrast to the first edition, the physical production of the book is excellent.

The main theme is the motion of electrons in metals, as manifested primarily in electrical conductivity. Conduction in metals, however, is dependent on so many factors that it can hardly be treated at all except as a culmination of a thorough study, such as is presented in this book, of properties and phenomena in which the various factors are in some degree isolated. After an introductory historical chapter, describing the development of the theory of conductivity from Drude to Sommerfeld, the motion of an electron in a perfect lattice is fully treated. Three chapters follow on metallic structures, alloys and semi-conductors. The "static" properties, thermal and magnetic, are fully discussed, and there is a shorter chapter on ferromagnetism. Conductivity and allied effects are treated in great detail in three chapters accounting for more than a third of the book, concerned with the formal theory, the mechanism of conductivity, and the application of the variation principle to conduction phenomena.

It is impossible, in a short notice, to do justice to the critical power and easy mastery of mathematical technique which are shown in this book. This is no stringing together of snippets hastily garnered from abstracts, but a comprehensive individual presentation of a complex branch of physical theory, built up from a careful study and working over of many difficult papers, among them the author's own. It would be idle to pretend that the book is easy to read, for, apart from the admirably written descriptive interludes, it is heavily mathematical. Indeed, some physicists may well feel that the powerful mathematical bombardment of the last twenty years or so has resulted in disappointingly small gains in the quantitative co-ordination of the stubborn detailed facts about the behaviour of actual metals, and the author would probably agree. For all those who are not content to deceive themselves with vain speculations, and wish to know what really has been achieved in the theory of metals or who wish to explore the field further this book will be indispensable. E. C. STONER

Reihenentwicklungen in der mathematischen Physik. By J. LENSE. (Berlin: Walter de Gruyter and Co.) Pp. 216. Price *DM*.26.

This is a mathematical text-book, written in German, containing six chapters which deal with the Euler–Maclaurin sum formula, the gamma function, orthogonal polynomials, Bessel functions, spherical harmonics, and Lamé functions. It presupposes a knowledge of contour integration. The mathematical reasoning is sound and careful, but the references given are more scholarly than helpful: whereas the earliest mention of an idea is traced to the papers of Euler, Bernoulli, etc., there are hardly any references either to modern or to English and American books. The needs of the physicist do not seem to have been sufficiently borne in mind: e.g. the Bessel function $K_n(x)$ is not mentioned, and the definitions of spherical harmonics differ from those which are now customary in quantum mechanics; and many similar examples could be quoted. Though the book contains good mathematical material in a compact form, it is likely to have only a limited appeal to physicists in this country.

J. M. C. SCOTT

Notes and comments

The Institute of Physics

At the Annual General Meeting of The Institute of Physics held on 18 May, 1954, Sir John Cockcroft (Director of the Atomic Energy Research Establishment at Harwell, and Chairman of the Defence Research Policy Committee) was elected President. Mr. G. R. Noakes was elected a Vice-President, Dr. S. Whitehead was re-elected Honorary Treasurer and Dr. B. P. Dudding re-elected Honorary Secretary. The two new Ordinary Members of the Board elected were Dr. K. A. G. Mendelssohn and Mr. H. P. Rooksby.

The Report for 1953 which was adopted at the meeting states that there was a net increase of 248 members in all grades and the total membership at the end of the year was 4595. The first holder of a Higher National Certificate in Applied Physics to become an Associate was elected during the year. Among the further qualifications recognized as complying with certain of the academic requirements of the membership regulations were the newly established degree in applied physics of the University of Manchester and the Associateship in applied physics of the Royal Technical College, Glasgow. The Report states that there is every indication that the Institute's Graduateship Examination, which was held for the first time in 1952, is meeting a definite need; it was conducted in four centres in 1953. The number of candidates for National Certificates in Applied Physics is increasing satisfactorily; there were 151 for the Ordinary Certificate and 55 for the Higher one.

The Institute's publications work continues to be one of its principal activities, the receipts from sales and advertisements amounting to £31 616 during the year. A notable feature of the publishing programme was the publication of the first four booklets (at 5*s*. each) in a new series entitled *Monographs for Students*; these are intended for general reading by students such as those in the first two years of a degree course in science or engineering or those reading for Higher National Certificates. They were well received and the sales were sufficiently encouraging for further booklets in the series to be commissioned. Dr. G. P. Barnard's *Modern Mass Spectrometry* was added to the *Physics in Industry* series. The circulation of both of the Institute's monthly periodicals, the *Journal of Scientific Instruments* and the *British Journal of Applied Physics*, continued to increase and the Report states that this *Journal* already has an established reputation.

The Institute's Education Committee is responsible for advising the Board on the many problems concerning education and training which arise from the rapid developments in pure and applied physics. Its suggestions and recommendations on questions which might involve either a change of policy on the part of official bodies (such as the universities or public examining bodies), or a change of outlook on the part of individual teachers in schools or universities were published in pamphlet form under the title *Problems concerning the Education of Physicists*.

The statistical results of a further enquiry among members about their remuneration were published during the year together with an analysis and discussion of this data. Like the previous enquiries on this subject (1948 and 1951) there was a considerable demand for copies from employers of physicists who find this data a valuable guide to them. The Report records that the permanent officers have again dealt with many enquiries from both members and non-members concerning professional matters.

The Report includes a brief account of the Institute's second four-day convention which was held at Bournemouth. These conventions are intended primarily for members and their ladies and some 450 took part in scientific meetings, social functions and outings. In addition to the several specialist conferences arranged by the Institute's subject groups, a three-day conference on Static Electrification was held in London in March, and was attended by some 200 persons. The proceedings were published as a Supplement to this *Journal*.

The Report shows that the Institute's sixteen local branches and specialist subject groups held many meetings during the year and visited various laboratories and works. Some of the groups also carried out other functions on behalf of their members.

Journal of Scientific Instruments

Contents of the July issue

SPECIAL ARTICLE

The Physical Society's Exhibition—London, 1954. By E. H. W. Banner.

ORIGINAL CONTRIBUTIONS

Papers

An ophthalmic magnet. By L. R. Blake.

A flexible single-channel pulse-amplitude analyser. By F. J. M. Farley.

A capacitance displacement gauge. By W. M. Todd.

An improved method of measuring dynamic elastic constants, using electrostatic drive and frequency-modulation detection. By H. Pursey and E. C. Pyatt.

A simple method of making vacuum-tight coolable window seals for low temperature optical transmission cells. By V. Roberts.

A stereoscopic reprojection apparatus for neutron scattering experiments. By M. H. Alston, A. V. Crewe and W. H. Evans.

An improved X-ray shadow projection microscope. By V. E. Cosslett and H. E. Pearson.

Laboratory and workshop notes

A water-separating cyclone. By A. M. Godridge and G. G. Thurlow.

A simple magnetically-controlled mercury cut-off. By V. J. Hammond.

A photoelectric flow rate meter. By W. R. S. Briggs and R. L. Werner.

A method for modifying the thermal e.m.f./temperature characteristics of constantan thermocouple wire. By W. R. Beakley.

An optical dip-stick for liquid air. By J. E. Geake.

Removal of water vapour and carbon dioxide from an infra-red spectrometer. By P. Holliday.

An easily operated fine voltage control. By H. House.

Note on the measurement of reflectivity and absorption of a transparent solid. By J. C. S. Richards.

NOTES AND NEWS

Correspondence

Suggested modification to the Unicam 9 cm X-ray powder camera. From R. L. Gordon.

A reversible aspirator for the thermal precipitator. From B. M. Wright.

A simple electronic relay for accurate temperature control. From R. S. N. Rau.

New books

New instruments, materials and tools

Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London.

It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone*: Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2*s*. 6*d*. including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone*: Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s*. each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

SPECIAL ARTICLE

Statistical concepts in theoretical physics*

By R. FÜRTH, D.Phil., F.Inst.P., F.R.S.E., Birkbeck College, University of London

The introduction of the principles of statistical physics into the curricula of students of physics at an early stage is advocated, and the need for explaining the apparent incompatibility of the concepts of causation and chance is emphasized. First the meaning of causal laws in physics is briefly discussed, and it is shown that it is not possible to prove the law of causation by experimental verification. Then the meaning of statistical laws is explained, and it is pointed out that they constitute true laws of nature which cannot be deduced by mere speculation. The opinion is expressed that statistical physics should be taught on an empirical basis rather than on the basis of "*a-priori* probability," the concept of which is critically examined. The conceptual difficulty, inherent in the kinetic theories, of reconciling the existence of macroscopic statistical laws with the presumed causal nature of the microscopic processes is pointed out, and it is indicated how this difficulty can be overcome by introducing the more general notion of "stochastic laws," the mathematical formalism of which is compared with that of the causal laws of classical physics. The establishment of stationary distributions and the meaning of *a-priori* probability are discussed from this point of view. Finally, it is suggested that the transition from classical to quantum physics can be made much easier for the students if they are first taught classical statistical physics in the indicated manner, as the laws of quantum mechanics can also be interpreted as stochastic laws.

The use of statistical methods in theoretical physics started with the development of the kinetic theory of gases at the beginning of the last century. Since then these methods have been applied on an ever-increasing scale, and as a result a whole new domain of theoretical physics, called "statistical physics" has emerged, with "statistical mechanics" and "statistical thermodynamics" as its most important branches. More recently, with the advent of quantum theory at the beginning of the present century, an even more radical change in the conceptual attitude of theoretical physics has taken place. The statistical interpretation of quantum theory, which stipulates that the laws of physics are essentially statistical in nature, has by now become almost generally accepted.

Consequently it has been found necessary to incorporate statistical physics as a separate subject in the curricula of students of physics and related branches of science at universities and similar institutions for higher studies and, to a certain extent, also in school curricula. But the way in which this subject is taught is not always satisfactory. The majority of students, who had been brought up on the Newtonian causal concept of the laws of nature, are suddenly and usually rather late in their courses brought in contact with the diametrically opposite concept of chance and probability of physical events, and this procedure tends to confuse them. In my opinion the students should be made acquainted with statistical ideas at an early stage in their courses of studies, and special emphasis should be laid on the teaching of the fundamental concepts involved and the explanation of the apparent incompatibility of the concepts of "causation" and "chance."

Many books and articles have been written on statistical concepts in physics, and I do not claim originality for most of what I want to say on this subject here. It is merely meant to give some help to teachers of physics who prefer making their pupils understand the fundamental concepts of science rather than supplying them with many details which could perhaps be learned better from textbooks.

Let us start with a brief discussion of the meaning of "causal laws" in physics and the way in which they must be formulated so as to be useful for the prediction of physical events. Two typical examples of such causal laws are the Newtonian equations of motion of material particles and Maxwell's equations for the electromagnetic field. The first requirement is that the law must be applicable to a great variety of objects, irrespective of their "individuality," for it would evidently be both useless and meaningless if it only held for one particular object and different objects obeyed different laws. For example, Newton's laws apply to all material particles, irrespective of their special properties, and Maxwell's laws apply to all kinds of electromagnetic fields. The second requirement is that the law should be independent of the special circumstances in which the object is placed; for example, the law of motion of a material particle must be formulated in such a way that it will govern the movement of the particle starting from any conceivable state and subject to any conceivable forces. This, of course, has the inevitable consequence that the trajectory of a material particle cannot be predicted from Newton's laws unless additional information concerning the initial position and velocity of the particle is provided, and the force acting on it is a given function of the particle's position and velocity for all time. Similarly the change in time of the electromagnetic field in a fixed volume of space can be predicted from Maxwell's equations only if the initial field is known, the "boundary conditions" are given for all time, and the electromagnetic parameters of the continuous media filling the space are known functions of space and time.

We have become so familiar with the working of cause and effect from personal everyday experience that we take the law of causation for granted. But when we attempt to formulate a causal law for a particular set of circumstances precisely in such a way that it can actually be used for the prediction of future events we find this very difficult, and we soon realize that the very existence of the law of causation is by no means evident. The question arises, therefore, whether it may be possible to prove the law of causation by experimental verification. How would one, for example, proceed in order to prove the existence of a causal law of motion of a material particle? One would have to make a great number of experiments in which one and the same test body was made to start its movement from the same initial position and with the same velocity under the same external conditions, and observe whether or not the body would arrive at the same final position after a given fixed time interval in every single experiment. In actual fact, the result of a set of such experiments will invariably be that the final states are *not* identical but spread out more or less into what is called a "distribution" of end states.

No experienced physicist will, of course, conclude from this that causal laws of motion do not exist. In many cases closer

* Lecture given on 14 April, 1954, in London, at the Annual Conference of the Education Group of The Institute of Physics.

inspection will show that the initial conditions were in fact not precisely the same in each single experiment. This is, for instance, the explanation for the apparently erratic behaviour of the balls in the well-known "Galton's sieve" experiment which is frequently used as a model for Brownian movement and for demonstrating how a "binomial" distribution of end states is generated (Fig. 1). In other cases the non-

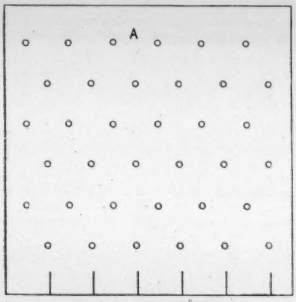


Fig. 1. Galton's sieve

A ball with a diameter slightly smaller than the distance between neighbouring pins is released at A and rolls down, threading its way between the pins in a zigzag path. The horizontal component of the movement of the ball is a one-dimensional "random walk" and can be regarded as a model for one-dimensional Brownian movement. If the experiment is repeated with a fairly large number of balls their distribution within the bottom compartments represents approximately a "binomial distribution."

reproducibility of a unique final state may be traced back to the insufficiency of maintaining constant external conditions (e.g. not sufficiently rigid mounting of the apparatus or insufficient thermal insulation).

But even in cases where no such obvious causes for apparent deviations from causality can be found it can still be maintained that there are certain "hidden" differences in the external or internal conditions of the test body in the individual experiments. In particular, apparently quite irregular fluctuations are superposed over the regular variations in time of an observed quantity; they are revealed by very sensitive observational instruments, like microscopes or amplifiers or counters, and which persist even under the most stringent preventive measures against disturbances. These fluctuations are generally taken to be the results of the irregular bombardment of the test body by the molecules of the surrounding medium, or of the uncontrollable internal movement of its constituent particles (e.g. Brownian movement of small particles, Brownian rotational movement of suspended systems, fluctuations of electric current, etc.) (Figs. 2, 3, 4). As it is impossible to reproduce experimental conditions precisely on a molecular scale the attempted proof for the causal nature of the laws of motion or other physical laws must fail, and our doubt concerning the existence of a rigorous law of causation in nature remains unresolved.

Indeed, when we go back to our everyday experience which led us to believe in cause and effect and on which the scientific notion of causation is based, we find that the number of cases in which the future course of events turned out to be quite unpredictable is very much greater than the number of instances in which a fairly accurate prediction could be made. It is therefore rather remarkable that the idea of universal causation has nevertheless gained such a strong foothold in the philosophy of science. To the naïve observer it must appear as if some events were the result of the working of cause and effect, but others simply due to "chance." Some people may be content with this state of affairs in which it is possible to forecast the outcome of certain actions, e.g. the lighting-up of a room when a switch is operated, but impossible to foresee certain other events, e.g. the result of a

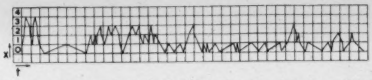


Fig. 2. Fluctuations in height of a particle in a liquid near the bottom of the vessel, due to the fluctuations of pressure on its surface. Instead of lying still at the bottom the particle constantly bobs up and down. (From an investigation on Brownian movement by R. Fürth.)

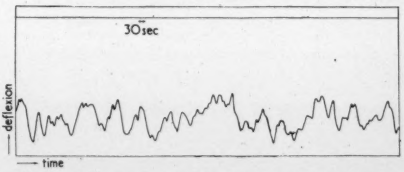
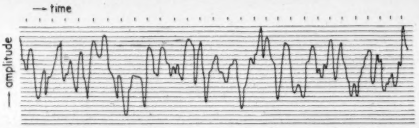


Fig. 3. Record of the deflexions of a supersensitive torsion balance showing irregular fluctuations in time due to the Brownian motion of the instrument. (From an investigation by E. Kappler.)



[Reproduced from *Proceedings of the Physical Society*]

Fig. 4. Traced record of the amplitude of electric fluctuations in the network of a wireless receiver. Frequency 100 kc/s. Timing marks 2 msec. (From an investigation on spontaneous electrical fluctuations by R. Fürth and D. K. C. MacDonald.)

bet. Other people, however, have an instinctive aversion against such a dualism of rules, by which the external world seems to be governed, and are convinced that they could predict even so-called "chance events" if they only knew the "laws of chance."

Although this seems at first sight a contradiction in terms, it is nevertheless a fact that certain limited forecasts can be made for seemingly random events, for example in chance games like throwing of dice or roulette. As a matter of fact, the search for rules for the successful playing of such games provided the impetus for the development of the mathematical theory of probability which started in the first half of the seventeenth century. We must therefore now try to analyse what is actually meant by "laws of chance."

If a certain phenomenon seems to be governed by chance and the result of an individual trial or the future development of an observed individual object cannot be predicted, the natural procedure seems to be to make a large number of

trials of the same kind or to make observations under uniform conditions on a large number of individual objects belonging to a defined "class." This is called a "statistical investigation." Its outcome is usually presented in the form of a list in which the relative numbers or "frequencies" of the observed results of the tests or observations are tabulated, irrespective of the individuality of the objects (for example, the relative number of times a certain combination of points is thrown with two dice, or the relative numbers of male and female births in a certain population during one year, etc.). The numerical material contained in this list is called a "statistical distribution" or colloquially a "statistics." If the same distribution is obtained every time the set of trials or observations is repeated, we shall be justified in maintaining that a "law of chance" exists or, more precisely, a "statistical law"; for it will then be possible to predict the statistical distribution of any future set of trials.

As in the case of the law of causation we have a strong instinctive feeling, based on personal experience, that a law of chance in this sense really holds; however, in order to prove it scientifically we must try to verify it by actually carrying out experiments of the kind indicated. But whenever this is done, even if the investigated material is very well defined and very strict conditions are imposed, the distribution of frequencies varies from one set to another to a certain extent, i.e. the results are *not* identical, and a precise prediction of the distribution for a future set is not possible. But it is also found that the deviations between distributions obtained for different sets of observations become smaller and smaller the larger the number of trials is made in each set, or briefly the larger the statistics is. It therefore seems plausible to assume that these deviations would vanish altogether if the statistics could be made infinitely large, and the presumed law expressing this belief is usually called the "law of great numbers." But it has to be kept in mind that just as we have shown that the existence of causal laws cannot be proved rigorously by experimental verification, the statistical law of great numbers must always remain hypothetical.

It is very important to impress upon students that statistical laws are true laws of nature and that they cannot be deduced by mere speculation, just as, for instance, the Newtonian laws of motion were not obtained as a result of a hypothesis alone without recourse to experience. The widely held belief that the law of great numbers can be derived mathematically from the calculus of probability is false. This latter branch of mathematics, which is logically consistent in itself, is only a method for calculating unknown "probabilities" from certain other known probabilities. If the calculus of probability is to be applied to actual physical phenomena these probabilities have to be related to the relative frequencies of distributions, and to make this relationship unique it must first be *assumed* that the law of great numbers holds.

Nevertheless it seems possible to predict the particular form of the distribution law under certain very special circumstances from considerations of symmetry alone, and then to derive more complicated distribution laws from these special ones, once their "*a priori* probabilities" have thus been determined, by the application of the probability calculus. One can, for example, easily visualize a model object of such symmetry that all its possible positions in a trial experiment are obtainable from each other by symmetry operations. In such a case the only logically conceivable distribution appears to be a uniform one, which means that equal *a priori* probabilities are attributed to each of the possible events. For instance, for a die of ideal cubic shape and ideal uniformity of material the *a priori* probability for any of the faces coming on top at any particular throw ought to be exactly 1/6. It is thus possible to derive, for example, the binomial distribution law for the combination of trials with several such dice.

So strong is our conviction of the stringency of symmetry considerations that discrepancies between an observed distribution and a theoretical distribution derived from *a priori* probabilities are generally taken as an indication that the observed phenomenon is not completely random in character. This argument is, however, open to grave objections. For in all physical phenomena we are dealing with *real* objects and not mathematical models, and as the former are never truly symmetrical the argument for the determination of *a priori* probabilities from symmetry considerations loses its logical basis. In fact a certain degree of asymmetry of the objects of observation is necessary for a statistical investigation to be made at all, as otherwise there would be no distinctive features on which to base a distribution; for example, the six faces of a completely symmetrical die would be undistinguishable. Consequently it is impossible to design any physical system which, when subjected to a statistical investigation, would yield a distribution that would not in at least some of its features show definite deviations from the expected ideal mathematical distribution. For instance, all so-called "random tables" of figures obtained empirically by means of statistical experiments on some mechanism have to be artificially adjusted so as to fit the requirements of randomness.

From all this it appears that the notion of *a priori* probability is of comparatively little value in statistical physics, and that the latter should rather be put on an empirical basis, and probabilities be defined as limits of relative frequencies of distributions. We have seen that in classical causal physics the task of the experimenter is to observe the *state* of a physical system and that of the theoretician to predict its future development from a tentative causal law, which can again be checked by observation. Similarly in statistical physics the observer determines a *statistical distribution*, and the theoretician tries to predict the change of this distribution in time on the basis of a tentative statistical law. This procedure will now be described in greater detail.

As already mentioned, statistical methods were first introduced in theoretical physics in connexion with the kinetic theory of gases. It soon became clear that the apparently causal form of the macroscopic gas laws is due only to the fact that the number of individual molecular events responsible for the outcome of a macroscopic observation is so extraordinarily large. For all macroscopically observable quantities are defined as averages over molecular distributions and thus, owing to the law of great numbers, precise predictions of the values of these quantities can be made. But the statistical character of the gas laws is directly demonstrated by the fact that *fluctuations* are observed when the sensitivity of the observational instruments is so much increased that the number of molecular events involved in an observation is greatly reduced. Thus, for example, fluctuations of density are revealed by a faint scattering of light in a gas, fluctuations of pressure by the Brownian movement of small particles suspended in a gas atmosphere, etc.

There remains, however, a great conceptual difficulty which is inherent in the classical kinetic theory of gases and which should not be skipped over lightly in the teaching of this subject: how is it possible to reconcile the existence of the statistical gas laws with the presumed causal nature of the laws governing the movement of the individual molecules of which the gas is made up? Some authors of textbooks and articles in this field seem to be of the opinion that it is the

lack of knowledge concerning the initial and external conditions of the individual molecular processes, which makes it possible to make precise statements about the behaviour of a multitude of molecules. But it surely is illogical to assume that the mere *lack* of some information could be responsible for an *increased* knowledge on some other aspects. Other authors again adopt a rather mystical attitude by introducing a hypothesis of "molecular chaos" into the theory as a counter-agent against strict causality. In my opinion these metaphysical speculations ought to be excluded from the teaching of statistical physics, and it should be emphasized that the basic principles can only be derived from experience by the process of induction; the procedure should therefore be the same as in the teaching of classical causal physics.

As already explained, the objects of statistical physics are not individual systems but multitudes or "assemblies" of similar systems, and these are quantitatively described by their distributions. It is maintained that general laws exist which make it possible to predict the distribution of an assembly at any future time, provided that its present distribution is known, and the "interaction forces" between the member systems and the "external forces" acting on these systems are given functions of time and space. It is seen that this formulation comprises also the previously considered "purely causal" and "pure chance" laws and does not require any further assumption about the separate existence of either the one or the other. It is, of course, not possible to *prove* the existence of these so-called "stochastic" laws (to use a modern term), just as we found it impossible to prove the existence of causal laws; the existence of stochastic laws remains hypothetical, and the only touchstone for this hypothesis is their successful application.

The information contained in a given distribution is clearly not sufficient for determining the fate of any of the members of the assembly; but this is not really a drawback as we are not interested in these individual happenings. It would be of little value if we were able to predict the shape of the trajectory of an individual molecule in a gas or of an individual electron in an electric conductor. The only entities of physical importance are in fact the distributions of assemblies because, as already remarked, all macroscopically observable physical quantities are defined as averages over such distributions. Indeed, the information given by the present distribution is just sufficient for calculating the values of these quantities at any future time. Mathematically this emerges from the fact that the mentioned stochastic laws have the form of partial differential equations (so-called "stochastic equations") for the distribution probabilities as functions of space and time, and their solutions are unique for given initial and boundary conditions. The complete analogy to the differential equations of causal physics, like the mechanical equations of motion and the equations of electrodynamics, is obvious.

A stochastic differential equation will in general have a solution which is independent of time; the corresponding distribution is called "stationary." There may, of course, exist more than one stationary solution, but if there is only one, it can evidently not depend on the initial distribution. Under these circumstances it is plausible to assume that an isolated assembly, when left to itself, will eventually, after a sufficiently long time has elapsed, settle down in this stationary distribution and remain in it indefinitely. We then say that the assembly is in "statistical equilibrium." The assumption that statistical equilibrium is set up automatically is a fundamental hypothesis of statistical mechanics and thermodynamics which essentially deal only with assemblies in statistical equilibrium. It is, however, not possible to prove this theorem on the basis of any "classical" consideration, and it is indeed doubtful whether it is generally true. It certainly does not hold exactly for finite assemblies of any kind, however large the number of their members may be; this is proved by the existence of the mentioned fluctuation phenomena which persist even after macroscopic equilibrium has been reached.

The just formulated theorem concerning the establishment of statistical equilibrium in infinitely large assemblies, if its correctness is accepted, constitutes one of the most important general laws of nature, comprising the theorem of the irreversibility of physical processes, as for instance expressed by the second law of thermodynamics, and also the law of great numbers. The distributions which are observed in the Galton sieve experiment and in chance games like dice and roulette are indeed simply stationary distributions of either real assemblies or ideal multitudes (sometimes called "statistical ensembles"), and we can now understand why these distributions are independent of the special "initial conditions," e.g. how a roulette ball is spun or dice are thrown. The so-called *a priori* probabilities are further seen to be the stationary solutions of particularly simple stochastic differential equations.

One of the greatest difficulties in the teaching of the fundamentals of quantum theory is to avoid a sudden break in the systematic build-up of the student's understanding of the concepts of physics when passing from classical mechanics to quantum mechanics and the statistical interpretation of quantum-mechanical equations. The indicated presentation of "classical" statistical physics may help in making this task much easier. For it is only necessary to show that, owing to the quantum nature of the mutual interaction of particles, it is not possible to observe the "state" of an individual system exactly, because the very process of measurement introduces a finite amount of uncontrollable disturbance. This makes it impossible, at least with our present means of observation, to test the law of causation even for isolated fundamental particles; modern physics therefore disclaims the law of causation for elementary processes by reason of the epistemological tenet that a statement is meaningless unless it can, in principle at least, be verified.

The determination of the distribution of an assembly of fundamental particles to any desired degree of accuracy is, however, not precluded by the quantum nature of interaction. This suggests that the laws of quantum mechanics have a stochastic character, that is that it ought to be possible to calculate these distributions at any time from the given initial distribution and, in particular, to determine the shape of stationary distributions which do not change in the course of time. Indeed, in the most widely used formalism of quantum mechanics the fundamental equations or "wave equations" have the form of partial differential equations for a "wave function" which can be solved for given initial and boundary conditions and permit of stationary solutions under certain conditions. According to the statistical interpretation of the wave equation the wave function is intimately connected with the "probability" of finding an individual system at a given time in a given configuration in ordinary or momentum space. From the previous discussion on the notion of probability it appears that these equations do not govern individual systems at all, but only assemblies of identical systems.

Thus the principles of quantum physics can be closely linked up with the principles of statistical physics, and I believe that after having attended a course built up on the lines indicated here, a student, instead of feeling bewildered, may actually experience a sense of mental satisfaction.

Sub-normal discharges

By Professor K. G. Emeleus, Ph.D., F.Inst.P., and J. A. Burns, M.Sc., A.Inst.P.,* Physics Department, The Queen's University of Belfast

[*Paper received* 15 *February*, 1954]

When very small currents are passed through mercury vapour in cylindrical tubes, the characteristic current *versus* voltage gradient curves for the positive column have negative slope. A previous analysis of conditions in the column based on ambipolar diffusion theory is restated and extended somewhat in terms of plasma-sheath theory. It appears that the plasma-sheath boundary region of high current discharges expands to fill the cross-section when the current is low. Probe experiments on positive columns in mercury vapour and other gases indicate the probable presence of positive space charges throughout the tube when the current density is small. It is suggested that this may be a common property of many discharges with small current density. References are given to literature on the subject.

1. Introduction

The theory of the ordinary uniform positive column in a cylindrical tube has been developed by Schottky[1] from the concepts of quasineutrality† and ambipolar diffusion, and by Tonks and Langmuir,[2] from the more fundamental concepts of the plasma and positive ion sheaths. Analyses of uniform positive columns by means of probes has shown reasonably good accord with theoretical predictions. It has also been shown that the properties of various modified unstriated forms of positive column can be understood qualitatively if the theory is extended to allow for radial inhomogeneity and cumulative ionization,[3] whilst other features of the positive column, in electro-negative gases, can be at least qualitatively understood from the properties of negative ions.[4]

The object of the present paper is to call attention to an investigation by Mierdel and Schmalenberg,[5] which appears to have escaped the notice which its importance warrants, on the uniform positive column in mercury vapour when the current density is very small, and to extend somewhat the work of these authors. Mierdel and Schmalenberg found that under these conditions there is marked increase of longitudinal field with decrease in current. For example, in a tube 5 cm in diameter containing mercury vapour at a pressure of 0·76 mm of mercury, the longitudinal field was found to remain close to 1 V/cm when the current density was reduced from 2×10^{-4} A/cm^2 to 2×10^{-6} A/cm^2, but rose with further decrease in current to 3 V/cm at a current density of 7×10^{-7} A/cm^2, and showed no sign of reaching a constant higher value. Similar results, with larger fields, were obtained at higher pressures. Curves showing the ratio of longitudinal field to pressure as a function of the product of tube radius and pressure at various currents moved upwards towards the sparking field curve, determined otherwise, as the current tended to zero.

Mierdel and Schmalenberg discussed these results in terms of Schottky's ambipolar diffusion theory, and pointed out that the condition of quasineutrality (effective absence of space-charge) will not be satisfied when the current density is very small. This condition is fundamental to the development of the normal theory, according to which the longitudinal field should be independent of current density if the density of the gas is constant.

An indication of how the failure arose was obtained as follows. Assuming that the concentration n_e of electrons falls off parabolically from a maximum value at the axis to zero at the wall ($r = R = 2{\cdot}5$ cm)—as an approximation to the law $n_e = n_0 J_0(2{\cdot}405\, r/R)$ of Schottky's theory—and that the radial variation of n_e with potential is governed by Boltzmann's relation for an electron temperature T, then to satisfy Poisson's equation there must be an excess concentration δn of positive ions at distance r from the axis given by

$$\delta n = \frac{kT}{\pi e^2} \frac{R^2}{(R^2 - r^2)^2} \qquad (1)$$

On the other hand, with the same parabolic law of variation for n_e, the electron density $n_e(r)$ at radius r will be given by

$$n_e(r) = \frac{2i(8mkT/\pi)^{\frac{1}{2}}}{\pi R^2 e^2 \lambda E}\left(1 - \frac{r^2}{R^2}\right) \qquad (2)$$

to ensure transport of the current i through the tube when the longitudinal field is E and the electron free path λ. The quasineutrality must certainly fail for r greater than the value S at which δn [equation (1)] and $n_e(r)$ [equation (2)] become equal. Inserting numerical values for the atomic constants, measuring E in V/cm, and the mean current density j in A/cm^2,

$$S = R[1 - 1{\cdot}91 \times 10^{-3}(E\lambda/j)^{\frac{1}{2}}T^{\frac{1}{4}}]^{\frac{1}{2}} \qquad (3)$$

It is remarkable how insensitive the expression for S appears to be to the value of T, and, to a lesser extent, to the values of E, λ and j, but it should be remembered that T and the product $E\lambda$ will increase or decrease together.

Mierdel and Schmalenberg showed from numerical consideration of equation (3) that the observed increase of E for very small currents almost certainly occurred when $R - S$ became comparable with R, i.e. when the condition for quasineutrality failed over a considerable part of the cross-section of the tube. They accounted for the increased longitudinal field under these conditions in general terms by supposing that the radial electric field near the walls became less than with larger currents, so permitting of more ready diffusion of electrons to the walls, which in turn necessitated a higher mean electron energy to maintain the ionization balance in the column.

Ecker[6] has recently succeeded in formulating the equations for a positive column without the restrictions imposed by quasineutrality, for the case of $\lambda \ll R$, which corresponds to the conditions of Mierdel and Schmalenberg's experiments, and has shown that the shapes of the current density *versus* longitudinal field characteristic curves found by approximate

* Now at General Electric Co. Ltd., Wembley.

† A quasineutral discharge is one in which the difference between the concentrations of positive ions and electrons is small compared with the concentration of either.

solution of his equations are in good agreement with Mierdel and Schmalenberg's experimental curves.

2. APPROACH FROM PLASMA-SHEATH THEORY

It is possible to arrive alternatively at what is essentially Mierdel and Schmalenberg's description of low-current discharges by starting from the plasma-sheath model of the positive column. An advantage of this approach is that it is not restricted to the "high-pressure" column ($\lambda \ll R$), as is the Schottky theory. For rough purposes, the ordinary positive column can be regarded as a system in which the larger part of the tube is occupied by a plasma, separated from the wall by a positive ion sheath. By far the greater part of the longitudinal drift current (arc current) which is equal to the circuit current outside the tube, is transported through the plasma, since the electron concentration diminishes rapidly on passing from the plasma into the sheath. The main effect of the sheath in this connexion is to reduce the effective conducting cross-section of the tube. As the arc current is diminished, the sheath expands towards the tube axis. Mierdel and Schmalenberg's theory of S [equation (3)] is practically a theory of the thickness of the wall sheath, defined in a particular way.

The simple plasma-sheath model of the positive column, if analysed in detail, is, however, known to lead to inconsistencies which can only be avoided by considering explicitly the boundary region between sheath and plasma.[2,7] The basic property of the boundary region is that it has a resultant positive space-charge which cannot be neglected, as is possible in considering most properties of the plasma, and an electron concentration which cannot be neglected, as is possible in considering most properties of the sheath. Since the transport of arc current requires the presence of electrons, and decrease of current leads, on the elementary model, to expansion of the wall sheath, the low-current form might, therefore, be regarded better as developed from the high-current form by expansion of the sheath-plasma boundary region until it occupied the whole cross-section of the tube. For intermediate value of arc current, it would still be possible to consider the axial part as nearly a plasma, and the peripheral part as nearly a sheath, but this model would become increasingly invalid as the arc current was reduced. The radial field in the boundary region is intermediate between the radial fields in sheath and plasma. Forcing more of the arc current to pass along the intermediate region would permit of more ready diffusion of electrons to the walls, and necessitate a higher longitudinal field, as in Mierdel and Schmalenberg's theory, provided that no overcompensating general increases in radial field were necessitated otherwise to maintain the discharge. A deeper analysis, such as Ecker's, is needed to show rigorously that the latter does not occur.

Without going to this detail, it is, nevertheless, possible to show, more explicitly than Mierdel and Schmalenberg have done, that for the case of $\lambda \ll R$, a species of positive column is likely to persist with very small current densities. The approximate theory can be developed by considering the radial movement of ions and electrons as due to diffusion, with an appropriate diffusion coefficient (D), the value of which depends on current density. For the larger currents, the value for D will be the ambipolar diffusion coefficient D_a of Schottky's theory of the normal column. As the current diminishes, the motion of the electrons will become increasingly governed by the ordinary electron diffusion coefficient D_e determined by collisions of electrons with neutral molecules, which may be two to three orders of magnitude greater than D_a.

The condition to be satisfied for maintenance of the column is

$$tI = 1 \tag{4}$$

where t is the average time spent by an electron in the free state in the gas, and I the average number of electrons produced by collision per second by an electron. The value of t will depend on the radial distribution of electrons in the tube, but cannot differ much from the Brownian expression $R^2/(aD)$, where a is a number of order of unity.

I is given by the relation

$$I = \int_{V_i}^{\infty} f(V) v P(V) dV \tag{5}$$

where $f(V)$ is the electron energy distribution function, V_i the ionization potential, v the velocity of an electron with energy eV electron volts and $P(V)$ the number of electrons produced by an electron of energy eV per cm of its random path. The mean energy eW of the electrons is likely to be of order of eV_i or less, so that $P(V)$ can be taken with sufficient accuracy to be a linear function of $V - V_i$. If $f(V)$ is Maxwellian, the integral becomes

$$AW^{\frac{1}{2}}(V_i + 4W/3) \exp\left[-3V_i/(2W)\right] \tag{6}$$

where A is proportional to the gas-pressure, but otherwise contains only atomic and numerical constants. The maintenance equation (4) can now be written as

$$\left[R^2A/(aD)\right]W^{\frac{1}{2}}(V_i + 4W/3) \exp\left[-3V_i/(2W)\right] = 1 \tag{7}$$

The diffusion coefficient can be written in the form

$$D = \lambda\bar{v}/(3g) \tag{8}$$

where λ is the mean free path of the electrons, $\bar{v}$ their mean velocity, and g depends on the density of ionization, ranging from unity when this is very small to D_e/D_a when it is so large that normal ambipolar diffusion is taking place.

Removing from equations (7) and (8) factors which contain only pressure, tube radius and atomic and numerical constants, the condition for equilibrium to be maintained, in the sense that if a uniform positive column discharge can be passed through a given tube at given pressure for one current it can also pass for another, is that for the complete range of values of g, W shall be capable of varying in such a way that

$$(1 + 4f/3) \exp\left[-3/(2f)\right] \tag{9}$$

can vary simultaneously by a factor $a\lambda/g$; f is the ratio W/V_i. Even allowing for a large Ramsauer effect, $a\lambda/g$ is not likely to alter by a factor of more than 10^4. The expression (9) can easily change under discharge conditions by a much larger amount. For example, since $V_i = 10{\cdot}4$ V for mercury, if $W = 0{\cdot}52$ V, it is $2{\cdot}5 \times 10^{-16}$, and if $W = 5{\cdot}2$ V, it is 8×10^{-2}.

For the smallest currents, the electron energy distribution may not be Maxwellian. In fact, the conditions for maintenance of a Maxwellian distribution through interaction of the charged particles in the gas may be much the same as those for application of quasineutral ambipolar diffusion theory, so that the analysis of conditions expressed by expression (9) may cease to be valid for just those small currents for which $D > D_a$ when the discussion is most needed. It is probable, nevertheless, that in any non-Maxwellian distribution actually occurring, a small change in mean energy, which will produce little change in D through change in $\bar{v}$ alone [equation (8)],

will produce a large enough change in the high energy tail of $f(V)$, and hence of I [equation (5)], to allow for the variation in g, which is essentially what has been shown for the Maxwellian distribution.

There is optical evidence that in other gases, under what appear to be generally similar conditions of low-current density, the positive column may be striated.[9,10] This will be accompanied by spatially periodic variations in the longitudinal field, the effect of which will probably be to increase the total ionization in the column above the value it would have for the same potential difference between its ends in a uniform column. The theory of the radial currents is, however, unlikely to be altered fundamentally.

3. EVIDENCE FOR POSITIVE SPACE-CHARGE

Irrespective of whether the properties of low-current positive columns are approached from ambipolar diffusion theory, or from plasma theory, it appears that a significant positive space-charge develops over the cross-section of the tube, as the current density is reduced. A probe test for positive space-charge has been applied by the authors to a mercury positive column passing a very small current. When a probe is put in a region of positive space-charge its characteristic current-voltage curve differs considerably from the normal Langmuir curve for a plasma.[11] Analysis of these curves for space-potential, etc., has not yet been achieved, because, with the small ionization density and pre-existing space-charge, the disturbance due to the probe may not be even approximately confined to its immediate neighbourhood, but will be more like that produced by a grid in an evacuated tube or thyratron passing a space-charge limited electron current. Their form is, however, readily recognized, and although they could conceivably also be produced for other reasons, they may be provisionally taken to show the presence of space-charge.

The discharge tube built for use with mercury vapour was 3·5 cm in diameter, with flat Nichrome disks perpendicular to the axis for anode and cathode, 16·6 cm apart. A molybdenum wire probe of diameter 0·1 mm and exposed length 3 mm projected radially from a side tube to about one third of the way across the tube, from a side arm, 2·85 cm from the anode. The probe was sheathed in thin glass, which did not touch it near the free end. The tube and a small bulb containing mercury attached to it were thermostated independently, and kept continuously exhausted by a mercury diffusion pump. The pressure of residual gases was less than 10^{-5} mm. Current was supplied by a d.c. generator. The probe circuit was arranged as usual, care being taken to avoid leak currents.

Probe characteristic curves obtained when very small tube currents were passing were of the expected unanalysed form. Typical conditions for their occurrence were, vapour pressure 0·65 mm, p.d. between anode and cathode 1240 V, tube current 0·8 μA; the tube shone with a faint green glow, concentrated somewhat axially, extending from the anode to about 4 cm from the cathode. Probe characteristics of the normal Langmuir type were not obtained until the current was increased considerably, e.g. to 16 μA at 2020 V and 0·88 mm pressure; under the latter conditions the appearance was that of a fully developed glow discharge. Between the steady low current and steady high current regimes, the tube had an intermediate state in which the discharge passed intermittently. Similar regions of instability have been found by Taylor[12] for cold cathode tubes with electrodes close together, and in this laboratory for a number of hot and cold cathode discharges. It is possible that certain distributions of space-charge in ionized gases are intrinsically unstable.

The authors' results for mercury agree with those of Greeves and Johnston for oxygen,[13] and of Emeleus, Greeves and Montgomery[9] for helium. [In all three gases part of the positive space-charge to which the bare metal end of the probes was exposed was that on its insulating stem. The relative extents to which the stems and tube walls can be regarded as the boundaries of the space-charges is necessarily uncertain, and would only be partly decided by an investigation like that made by Sloane and Emeleus[14] for plasmas. It is important to remember that, even in a plasma; (1) the effect of the space-charge on the stem may make data obtained with a small exposed piece of metal less reliable than those obtained with a rather larger exposed piece, and (2) some disturbance of the discharge is produced by the insulating probe stem, whatever the size and potential of the exposed metal part of the probe.] Positive columns similar in appearance had been noticed previously when very small currents were passed through argon, but no attempt was then made to analyse them.[15] They can probably be produced in at least any electro-positive gas; it is not clear what effect the tendency of negative ions to accumulate axially might have in a strongly electro-negative gas when the current density is very low.[16] [In spite of the existence of stable negative ions of oxygen, discharges through this gas do not necessarily show the characteristic properties of discharges through strongly electro-negative gases (*cf.* Seeliger and Wichmann[17]).]

4. GENERAL CONSIDERATIONS

The d.c. positive columns considered in Sections 1–3 belong to a group of discharges, or sections of discharges, of which comparatively little is known; in all, the concentration of charged particles is insufficient to give a plasma and a positive sheath on the walls or cathode. It is suggested that they may be termed "sub-normal". Other examples are:

(i) the so-called low-pressure "corona" discharges between large electrodes held fairly close together, investigated by Seeliger and Schmekel,[18] Taylor,[12] and others.[10] The current *versus* voltage characteristic curve has usually a negative slope, like a sub-normal positive column. The appearance of these discharges approaches that of the negative sections of ordinary glow discharges as the current increases;

(ii) low-current discharges occurring during the period of formative lag of higher-current glow discharges;[19]

(iii) discharges in hot cathode tubes such as thyratrons and gas-filled rectifiers when the applied voltage is greater than the ionization potential of the gas, but an arc-type discharge has not struck;

(iv) the aureoles surrounding the cores of constricted positive columns.[3,20]

If the energy supply to the discharge is restricted, a plasma-type discharge may never develop from the subnormal discharge when this would otherwise be possible. If the supply is less restricted the sub-normal characteristic curve may be the threshold curve for initiation of a plasma-type glow discharge.[12,21] The dynamic and static subnormal characteristic curves will usually be not quite the same.

The discharges fall roughly into two classes, according to whether the charged particles produced in the gas move mainly to the electrodes (particularly the cathode), or the walls. Long tubes in which wall diffusion is predominant must naturally have terminal sections in which the electrode

losses characteristic of short tubes predominate. The distinction between the two classes is not, however, sharp, and it is conceivable that decrease in current in a subnormal positive column could raise the longitudinal field in it to that in the subnormal cathode section of the same discharge. With present knowledge of the theory of these dis-discharges,[6,10,22,23] it could not, however, be readily predicted in any particular case if this would occur before the discharge was extinguished, or, if it occurred, if the transition would proceed smoothly or discontinuously.

It is probable that the space-charges in subnormal discharges are positive, except where the flow of current is controlled by electron space-charge limited emission from a hot cathode or gas cathode.

ACKNOWLEDGEMENTS

We wish to thank Miss J. W. Beck for help with the experiments, and Dr. G. Ecker and Mr. G. J. McEwen for their comments on the paper.

REFERENCES

(1) SCHOTTKY, W. *Phys. Z.*, **25**, p. 635 (1924).
(2) TONKS, L., and LANGMUIR, I. *Phys. Rev.*, **34**, p. 876 (1929).
(3) SEELIGER, R. *Phys. Z.*, **33**, pp. 273, 313 (1932).
(4) EMELEUS, K. G., and SAYERS, J. *Proc. Roy. Irish Acad. A*, **44**, p. 87 (1938).
(5) MIERDEL, G., and SCHMALENBERG, W. *Wiss. Veröff. Siemens-Konz.*, **15**, p. 60 (1936).
(6) ECKER, G. *Proc. Phys. Soc.* [*London*], *B*, **67**, p. 485 (1954).
(7) BOYD, R. L. F. *Proc: Roy. Soc. A*, **201**, p. 329 (1950).
(8) cf. EMELEUS, K. G., and DEAS, H. *Phil. Mag.*, **40**, p. 460 (1949).
(9) EMELEUS, K. G., GREEVES, F. D., and MONTGOMERY, E. *Proc. Roy. Irish Acad. A*, **43**, p. 35 (1936).
(10) DRUYVESTEYN, M. J., and PENNING, F. M. *Rev. Mod. Phys.*, **12**, p. 87 (1940).
(11) EMELEUS, K. G., and BROWN, W. L. *Phil. Mag.*, **22**, p. 898 (1936).
(12) TAYLOR, J. Thesis, Utrecht (1927).
(13) GREEVES, F. D., and JOHNSTON, J. E. McF. *Phil. Mag.*, **21**, p. 659 (1936).
(14) EMELEUS, K. G., and SLOANE, R. H. *Phys. Rev.*, **44**, p. 333 (1933).
(15) EMELEUS, K. G., and HARRIS, N. L. *Phil. Mag.*, **4**, p. 49 (1927).
(16) cf. SEELIGER, R. *Ann. Phys.* [*Leipzig*], **6**, p. 93 (1949).
(17) SEELIGER, R., and WICHMANN, A. *Ann. Phys.* [*Leipzig*], **9**, p. 235 (1951).
(18) SEELIGER, R., and SCHMEKEL, J. *Phys. Zeits.*, **26**, p. 471 (1925).
(19) PENNING, F. M. *Phys. Z.*, **27**, p. 187 (1926).
(20) EMELEUS, K. G., SLOANE, R. H., and CATHCART, E. B. *Proc. Phys. Soc.* [*London*], **51**, p. 978 (1939).
(21) APPLETON, E. V., EMELEUS, K. G., and BARNETT, M. A. F. *Proc. Cambridge Phil. Soc.*, **22**, p. 447 (1925).
(22) LOEB, L. B. *Fundamental Processes of Electrical Discharge in Gases* (New York: J. Wiley and Sons Inc., 1939).
(23) ALLIS, W. P., and ROSE, D. J. *Phys. Rev.*, **93**, p. 84 (1954).

Triaxial stresses caused by notches

By A. F. C. BROWN, B.Sc.(Eng.), A.M.I.Mech.E., National Physical Laboratory, Teddington, Middlesex

[*Paper first received* 11 *March and in final form* 9 *April*, 1954]

The purpose of the work described was to give metallurgists a guide on means of suppressing plastic deformation in ductile materials, so as to provide a test for brittleness. Stress analyses were made on round bars with circumferential notches by the photoelastic frozen stress method and mechanical tests on steel test pieces provided additional information. An important conclusion is that, from the point of view of suppressing general yielding, the stress distribution at the middle of the bar has much more effect than that at the root of the notch. There is a need for further work on other shapes of notch.

INTRODUCTION

In a round bar in tension with a sharp circumferential notch, general plastic deformation occurs at a higher value of the mean shear stress across the reduced section than in a plain bar, because the radial and circumferential stresses set up at this section reduce the average effective shear stress. Therefore, by suitably notching a test piece, the onset of plastic deformation may be delayed and the possibility of achieving brittle failure may be increased. In order to exploit this means of extending the knowledge of the brittle strength of materials to the full, metallurgists urgently need more precise information on the stress distribution under the notch and more particularly they need to know how to compute the value of the average effective shear stress.

Theoretical solutions for the stress distribution only exist for very deep and very shallow notches. Neuber's solution for an infinitely deep notch of hyperbolic form[1] implies a test piece of infinite diameter. At the other extreme there are analytical solutions for very shallow notches covering quite a wide range of notch forms. There is therefore a need for precise information on notches of moderate depth. This information can be obtained by the photoelastic frozen stress technique but the method is slow and there are practical limitations to the forms and depths of notch which can be cut and loaded with sufficient accuracy. Moreover, in applying the results of photoelastic investigations to ordinary engineering materials, allowance must be made for the effect of the considerable difference in Poisson's ratio between the photoelastic material and the metal which it represents. For this assessment there is little guidance beyond that to be derived from the effect of Poisson's ratio in the two limiting theoretical cases.

Even if the stress distribution is known, it is still not easy to infer the condition for general plastic deformation. Immediately under the root of the notch, the axial stress is high but the radial stress is zero so that, even though there may be considerable circumferential stress, the local maximum shear stress is half the tensile stress and the material in this small region may be expected to yield just as it would in a plain test piece. On the other hand, this local yielding may

not affect the stress distribution at regions far away from the notch root and general yield will not ensue until the average effective shear stress reaches a value comparable with that which would cause yield in a plain test piece. It is reasonable, therefore, to postulate that the material always yields under the same value of the average effective shear stress and to take this postulation as defining the meaning of "average effective shear stress." Then by comparison with the stress distribution, in such few cases as it may be known, the manner in which the "average effective shear stress" should be computed from the stress distribution may be assessed. Thus, for any form and depth of notch to which such analysis may be applicable, the ratio of the average tensile stress in the axial direction across the notch plane to *twice* the corresponding average effective shear stress may be designated the triaxiality ratio. From examination of the variation of this ratio with respect to notch form and notch depth, the conditions most favourable to the production of brittle failure may be indicated.

In an endeavour to develop these conceptions quantitatively, complete determinations have been made by the photoelastic frozen stress technique of the stress distributions under two forms of notch and it is intended eventually to examine one or two more. However, at the present time the major part of the evidence regarding triaxiality ratio has been derived from the results of actual tensile tests on pieces of mild steel notched to various depths with notches of two different forms. Accordingly, the evidence of these tests will be presented first.

STEEL CYLINDERS WITH CIRCUMFERENTIAL V-NOTCHES IN TENSION

The steel on which the stress determinations were made was used in the normalized condition and had the following composition:

carbon	0·14 to 0·15%
silicon	0·22 %
manganese	0·82%
sulphur	0·049%
phosphorus	0·030%

The notches chosen for investigation were 45° V-notches and Fig. 1 shows the range of sizes covered. Starting with a plain bar and continuing with notches of increasing depth, extensions were measured with Marten's extensometers on a 2 in. gauge length spanning the notch. Stress/strain curves were plotted and from these the 0·1, 0·5 and 1% proof stresses were deduced, i.e., the stresses required to produce 0·1, 0·5 and 1% permanent strain. In computing the strain

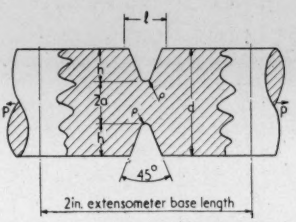


Fig. 1. Proportions of notches studied

Ranges covered { d = 1 in. and $\frac{1}{2}$ in.; $2h/d$ = 0·2 to 0·9; ρ/d = 0·01 and 0·05.

on a notched test piece, use was made of an effective gauge length l (Fig. 1). This length is that of a cylinder of radius a, which can be imagined as replacing the notch and giving the same elastic extension under load. The values of l, which have been calculated from the initial slopes of the stress/strain diagrams are given in Table 1, together with other measured and derived quantities.

Typical stress/strain diagrams in Fig. 2 show that the 1% proof stress is not very sensitive to the exact value of l, owing to the rapid increase of strain which occurs at the initiation of yield. The 1% proof stress has therefore been used to calculate the triaxiality ratio n, expressed as the proof stress on a notched test piece divided by the proof stress on a plain test piece, the latter being twice the average effective shear stress. The values of n, which are listed in the last column of Table 1, are shown plotted in Fig. 3. With the exception of that for the test on the very deep notch ($2h/d = 0·9$), where the test piece may have been cracked in machining, all the results show a steady increase of n as

Table 1. *Results of tensile tests on circumferentially notched steel cylinders*

d* (in.)	ρ (in.)	h (in.)	$\frac{2h}{d}$	$\frac{2a}{d}$	$\frac{a}{\rho}$	l (in.)	Proof stress (tons/in²) 0·1%	0·5%	1·0%	n = 1% proof stress on notched specimen / 1% proof stress on plain specimen
0·993	∞	0·0	0·0	1·0	0·0	—	20·4	20·4	20·4	1·0
1·000	0·050	0·100	0·20	0·80	8·0	0·25	23·7	24·4	24·4	1·20
1·000	0·050	0·200	0·40	0·60	6·0	0·41	28·2	31·0	31·0	1·52
0·999	0·050	0·250	0·50	0·50	5·0	0·41	29·0	32·6	33·1	1·62
1·000	0·046	0·352	0·70	0·30	3·0	0·34	35·3	42·9	46·5	2·28
0·989	0·050	0·393	0·79	0·21	2·0	0·31	40·8	45·6	47·1	2·31
0·990	0·010	0·10	0·21	0·79	39·5	0·31	22·5	24·7	25·0	1·23
0·999	0·010	0·200	0·40	0·60	30·0	0·38	28·0	32·2	33·8	1·66
0·998	0·010	0·249	0·50	0·50	25·0	0·39	29·8	35·1	37·4	1·83
1·000	0·009	0·349	0·70	0·30	16·7	0·32	35·2	41·5	44·4	2·18
0·997	0·009	0·410	0·82	0·18	9·8	0·25	43·2	49·0	50·6	2·48
0·940	0·012	0·422	0·90	0·10	4·0	0·12	37·6	41·4	44·7	2·19
0·502	0·025	0·099	0·39	0·61	6·1	0·22	24·7	31·7	32·6	1·60
0·492	0·025	0·123	0·50	0·50	5·0	0·17	26·7	33·0	34·4	1·69
0·500	0·025	0·149	0·60	0·40	4·0	0·24	30·7	36·2	38·7	1·90

* $d = 2(a + h)$

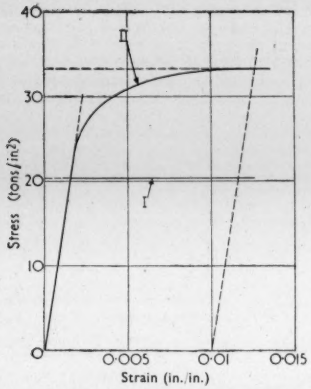


Fig. 2. 1% proof stress in tension on plain cylinder and typical notched cylinder

Curve I: plain cylinder.
Curve II: circumferential notch $\begin{cases} 2h/d = 0{\cdot}5 \\ \rho/d = 0{\cdot}05 \end{cases}$

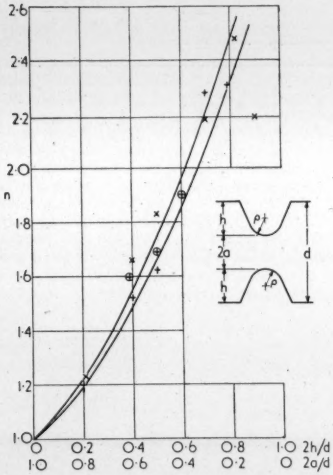


Fig. 3. Circumferentially notched steel cylinders in tension

Triaxiality factor, n, expressed as

$$\frac{\text{1\% proof stress on notched specimen}}{\text{1\% proof stress on plain specimen}}$$

d (in.)	ρ/d = 0·01	ρ/d = 0·05
1·0	×	+
0·5	–	⊕

the notch deepens. In contrast with the large effect which the root radius of a notch has on the more familiar stress concentration factor, its effect on the triaxiality ratio is small, notch depth being the dominating factor. The fact that high triaxiality ratios are associated with deep notches strongly suggests that the most critical form of test for brittleness will be one where the notch is as deep as is consistent with convenience in practice.

PHOTOELASTIC STRESS ANALYSES ON CIRCUMFERENTIAL V-NOTCHES

The two three-dimensional stress analyses by the photoelastic frozen stress method were also made on 45° V-notches in tension. Fosterite models of the same size as the steel specimens were used and longitudinal and transverse slices were cut. The retardations were measured by the Tardy method, using the N.P.L. photoelectric photometer for indicating extinction.[2] The curvatures of the stress trajectories were measured directly with the same equipment[3] and the principal stresses at the notch plane were separated by means of Jessop's adaptation of the Lamé–Maxwell equations for planes of symmetry.[4]

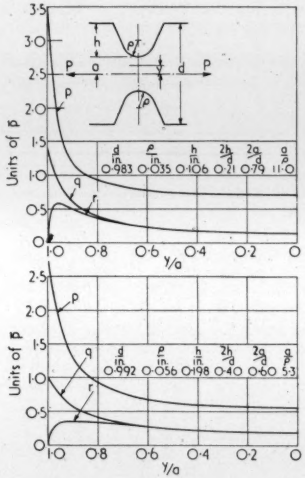


Fig. 4. Photoelastic stress analyses on circumferentially notched cylinders in tension. Principal stresses at notch plane

p = axial stress.
q = circumferential stress.
r = radial stress.
$\bar{p} = P/\pi a^2$.

The results are given in Fig. 4, in which the three principal stresses acting in the plane of the notch are expressed in terms of the average axial stress and shown as a function of the radius. The general form of the curves is as would be expected, but it is unfortunate that time has not permitted the examination of other cases. The two notches have approximately the same ratio of depth to root radius so, as would be expected, the shallower one has the higher stress

concentration factor, the values of the two notches being 3·5 and 2·7. The axial stress is also relatively higher for the shallower notch at the middle of the bar but slightly lower just inside the notch. The deeper notch causes larger circumferential and radial stresses at the middle and this affects the triaxiality ratios $[n = p/(p - r)]$, shown at the bottom of Fig. 5.

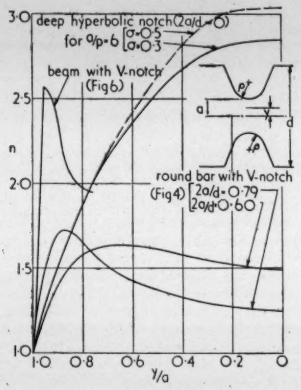


Fig. 5. Triaxiality ratio (n) at different positions in notch plane

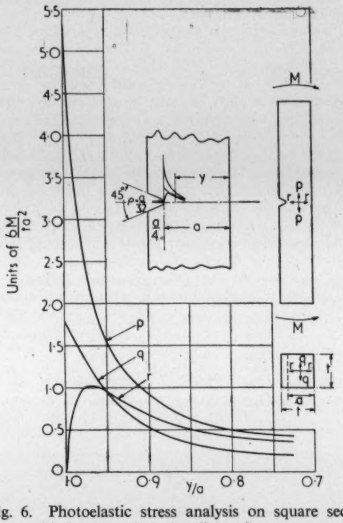


Fig. 6. Photoelastic stress analysis on square section beam with Izod notch in bending. Stresses on axis of symmetry

PHOTOELASTIC STRESS ANALYSIS ON A BEAM IMPACT SPECIMEN WITH A V-NOTCH

In view of the great use which is made of the notched-bar impact test it was thought that a similar stress analysis on this form of test piece would be valuable. A square beam with a transverse V-notch of the same form as that used in the Izod test was chosen for investigation and Fosterite models 2 in. square by 12 in. were made, the large size being dictated by the small root radius of the notch. Two models were loaded simultaneously, as in this case there was no axial symmetry; the stresses were determined along a line perpendicular to the centre of the notch and two slices intersecting at this line were needed.

The results of this analysis are given in Fig. 6, the stresses being expressed as multiples of the bending stress in an unnotched bar having a depth equal to the depth under the notch. The three-dimensional nature of the stress system is clearly shown, the stress along the notch, q, being one-third of the bending stress, p, at the notch root. The triaxiality ratio, given in Fig. 5, is high, the discontinuity in the graph being caused by the crossing of the q and r curves at a small depth below the notch. This curve is not strictly comparable with the others in the same figure owing to the relatively small root radius of the notch.

THE EFFECT OF POISSON'S RATIO ON THE TRIAXIALITY RATIO IN THE CASE OF A DEEP HYPERBOLIC NOTCH

A difficulty arises in comparing the results of the several forms of test from the fact that stress distributions are dependent on Poisson's ratio, σ, which is about 0·3 for steel, but nearer to 0·5 in Fosterite. Some light can be thrown on its effect from a consideration of Neuber's theoretical solution

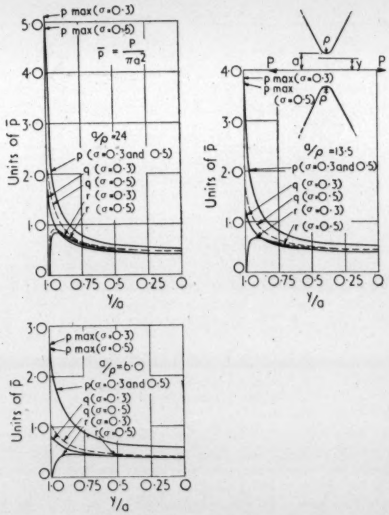


Fig. 7. Neuber's solution for deep hyperbolic notches in tension

p = axial stress.
q = circumferential stress.
r = radial stress.

for the infinitely deep hyperbolic notch[1] into which σ enters. The stress distributions for three notches as defined by the factor a/ρ have been evaluated and the results are given in Fig. 7. The axial stress, p, is little affected by changing σ from 0·3 to 0·5 but the circumferential stress, q, is altered considerably, particularly when the notch is sharp (a/ρ large); r also is affected to some extent. The triaxiality ratio for the case of $a/\rho = 6$, where the asymptotes to the hyperbolae are at approximately 45°, has been plotted in Fig. 5 and the effect of changing σ is seen to increase n by 7% at the middle of the bar. This difference for the sharper notches is greater. It is difficult to draw general conclusions on the effect of σ from Neuber's curves, because it is not possible to consider cases, for instance, where the angle of the notch is fixed and the root radius varied. It is probably safe to conclude, however, that for the two circumferential notches treated photoelastically, the effect of σ is not important, but that in the case of the beam test piece, it might be considerable.

CORRELATION OF DATA

An attempt is made in Fig. 8 to correlate the results on the circumferential V-notches in tension. The triaxiality

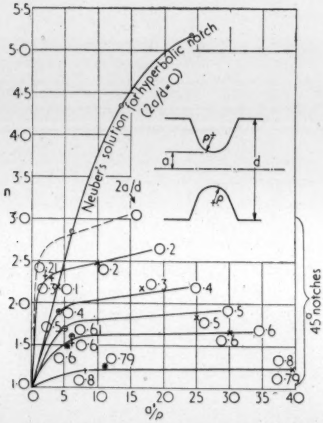


Fig. 8. Triaxiality ratio (n) for circumferential V-notches in tension. Figures against points are values of $2a/d$

+, ⊕, × } tests on steel cylinders from Fig. 3.
* photoelastic tests from Fig. 4.
○ Neuber's theoretical solution.

ratio, n, is shown on a basis of a/ρ, the relationship between the width at the notch plane and the root radius. The different curves represent various notch depths, the dotted curve for $2a/d = 0$ being sketched in to indicate the case of a deep 45° notch. In the construction of the diagram the results obtained from the proof stresses on mild steel specimens were entered first. It was then found that, for the two photoelastic test results, only the value of n at the middle of the bar conformed with those on the steel specimens. This is clearly shown in Table 2.

(The maximum values are from the peaks of the curves in Fig. 5, the average values are worked out on a cross-sectional area basis and the middle values are those for $y/a = 0$.)

Whether initial yielding occurs at the notch root or not, a comparison of the figures in Table 2 with the values for the steel specimens in Fig. 8 does suggest that it is the stress distribution at the middle of the specimen which is the important factor in the initiation of general yield. This finding seems to be an important one and suggests that the triaxiality ratio may be more important than the stress concentration factor under certain conditions of loading with ductile materials.

Table 2. *Triaxiality ratio from photoelastic tests of Fig.* 3

$2a/d$	n maximum	n average	n middle
0·79	1·72	1·48	1·25
0·60	1·63	1·51	1·49

The theoretical values of n for deep hyperbolic notches are based again on middle values and for a Poisson's ratio of 0·3. The curve joining them cuts across the curves for the 45° notches because the effective angle of the hyperbolic notch becomes large at low values of a/ρ. The dotted curve has been made to cut the theoretical curve at the point where the angle between the asymptotes of the hyperbola is 45°. A similar family of curves could be constructed for some other notch angle, say 30°, when the value of a/ρ for a deep notch would be 13·9. At this value of a/ρ, the triaxiality ratio for a 30° notch of infinite depth is 4·1, which is nearly 40% greater than that for a 45° notch. This suggests that it might be more convenient to use a 30° notch of moderate depth rather than a deep one with an included angle of 45°.

By turning again to the curves of Fig. 5, the effect of notch depth on the triaxiality ratio can be seen from a comparison of the curves for the deep hyperbolic notch ($2a/d = 0$) with the two curves obtained from photoelastic measurements ($2a/d = 0·6$ and $2a/d = 0·79$). With a deep notch, the value of n is high at the middle of the bar but, as the notch becomes shallower, so n becomes less at the middle. At the same time there is a tendency for n to rise to a maximum near the root of the notch. From these facts the shape of the curve for the very shallow notch can be inferred. Here the disturbance due to the notch must be expected to be extremely local with a high value of n near the root. At the middle of the bar the effect of notching will not be felt at all and the triaxiality ratio will fall to unity.

ACKNOWLEDGEMENTS

The work described was carried out as part of the research programme of the National Physical Laboratory and the paper is published by permission of the Director. The author wishes to acknowledge the assistance and co-operation of Mr. H. L. Cox, and the interest of Dr. N. P. Allen of the Metallurgy Division, who initiated the work.

REFERENCES

(1) NEUBER, H. (Berlin: Julius Springer, 1937.) David Taylor Model Basin Translation No. 74. *Theory of Notch Stresses: Principles for Exact Stress Calculations.*
(2) BROWN, A. F. C., and HICKSON, V. M. *Brit. J. Appl. Phys.*, **1**, p. 39 (1950).
(3) BROWN, A. F. C. *Brit. J. Appl. Phys.*, **2**, p. 138 (1951).
(4) JESSOP, H. T. *J. Sci. Instrum.*, **26**, p. 27 (1949).

The effect of latent heat on numerical solutions of the heat flow equation

By P. H. Price, B.Sc., Ph.D., and M. R. Slack, B.Eng., Ph. D., Department of Fuel Technology, University of Sheffield

(*Paper received* 2 *November*, 1953).

The stability and accuracy of finite difference solutions of the heat flow equation which involve latent heat evolution are discussed. A comparison of true and finite difference solutions is used to show how a numerical solution must be interpreted, and a dimensionless group is developed which governs the appearance of inaccuracies peculiar to numerical solutions involving a latent heat term.

A previous paper by the present writers[1] discussed the stability and accuracy of some numerical methods of solving the heat flow equation in cases where the thermal diffusivity is a function of temperature. Further investigation is required when a latent heat effect also occurs. It has been found convenient to examine two types of latent heat effect.

The first type is the case of a small latent heat effect distributed over a range of temperature as in the case of the change point in steel. The second type consists of a large latent heat effect occurring isothermally, as in some cases of solidification or evaporation.

In the first of these cases the latent heat effect can be added to the specific heat, reducing the problem to one involving variable diffusivity, which was discussed in the previous paper. The stability condition then involves the rate of change of diffusivity with temperature.

When the latent heat is of the second type the rate of change of diffusivity becomes infinite and the stability analysis is no longer applicable. This paper is therefore devoted to a discussion of this case of a latent heat occurring at a fixed temperature.

For simplicity of exposition, the case of heat flow in a one-dimensional slab is considered. Of the following discussion, the conclusions on stability may easily be generalized to include two or three dimensions, but to extend the investigation of accuracy would be more difficult.

THE EFFECT OF LATENT HEAT ON STABILITY

The characteristic feature of a numerical solution in which a latent heat is evolved, is that each mesh point in turn reaches and remains at the temperature at which the latent heat effect occurs, until all the latent heat at the point is developed. Therefore, whilst any one mesh point remains at the constant latent heat temperature, the slab could be partitioned at that point, and boundary conditions of constant surface temperature applied to the two new surfaces formed. The stability of the solutions in each of the smaller slabs formed may be investigated by the methods already known,[1-3] and no new problem arises.

THE EFFECT OF LATENT HEAT ON ACCURACY

Latent heat is evolved at one mesh point during the time in which its temperature moves right across the lamina which the mesh point governs. The temperatures at that mesh point as given by the numerical solution will therefore be in error until the phase change front is some way past the lamina, when the numerical and true solutions should approach each other again. Furthermore, it is possible for the front to remain at one mesh point so long that the constant temperatures which pertain at the phase change eventually cause the calculated temperature distribution in the conductor to reach the steady state.

Numerical example. The inaccuracies which can arise as a result of the method of representing latent heat evolution are illustrated by the following example for the case of a semi-infinite volume of water initially at 0° C. The surface temperature is suddenly lowered to −50° C and maintained at that temperature. The resulting temperature distribution is governed by the equation:—

$$\rho \frac{\partial H}{\partial t} = k \frac{\partial^2 \theta}{\partial x^2} \qquad (1)$$

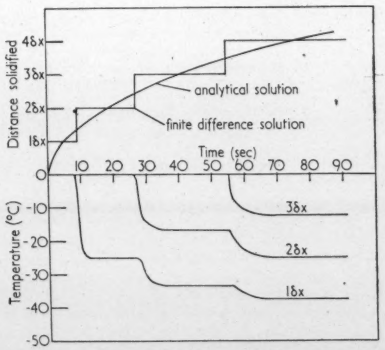
where H = heat content per unit mass;
ρ = density;
θ = temperature;
t = time;
x = distance;
k = thermal conductivity.

A finite difference form of equation (1) is

$$\rho H_{m,r+1} = \rho H_{m,r} + (\theta_{m+1,r} + \theta_{m-1,r} - 2\theta_{m,r})k \,.\, \delta t/(\delta x)^2 \qquad (2)$$

where $x = m \,.\, \delta x$, $t = r \,.\, \delta t$.

The numerical solution is obtained by applying equation (2) successively to each point in the mesh; the temperature distributions at time t being used to find heat content distributions at time $t + \delta t$. The relation between heat



Upper curves: progress of the solidification front, as predicted by the numerical and analytical solutions

Lower curves: temperatures at the first three mesh points, as given by the numerical solution

content and temperature is known from the thermal properties, which for ice are:

specific heat $c\rho = 0{\cdot}461$ cal. c.c.$^{-1}$ °C^{-1}
thermal conductivity $k = 0{\cdot}0053$ cal. cm^{-1}. sec^{-1}. °C^{-1}
latent heat $L\rho = 78{\cdot}6$ cal/c.c.

A graph of $H\rho$ *versus* θ was constructed before the calculation was commenced. δt was chosen as 1 sec and stability ensured by taking $\delta t = 0{\cdot}8(\delta x)^2/2D$ as discussed in the previous paper.[1]

The values of the temperatures which result at $2.\delta x$ and $3.\delta x$ inside the slab are shown in the figure, together with the progress of the solidification front as given by the finite difference calculation. The manner in which the steady state is reached at mesh points in the ice is clearly shown. Neumann's analytical solution of the differential equation, reported by Carslaw and Jaeger,[4] for the progress of the solidification front is also included, and this is straddled by the stepwise progress calculated in the numerical solution.

Comparison of true and numerical solutions. It appears from the results of the above example quoted in the table, that the time where the solidification plane is actually passing a mesh point occurs when the numerical solution records that half the latent heat is lost.

Comparison of solidification times as given by the analytical solution and the time at which half the latent heat is lost in a numerical solution

Distance solidified	*Time for solidification as given by Neumann's analytical solution (sec)*	*Time in numerical solution when half latent heat at mesh point is lost (sec)*
$1.\delta x = 0{\cdot}1696$ cm	4·64	4·26
$2.\delta x = 0{\cdot}3392$ cm	18·5	18·4
$3.\delta x = 0{\cdot}5088$ cm	41·7	41·7
$4.\delta x = 0{\cdot}6784$ cm	74·2	74·6

The agreement between numerical and true solutions is analysed further as follows.

Consider the problem in which the region $x > 0$ contains liquid at its freezing point T, of latent heat L, and let the surface temperature be instantaneously reduced to $\theta = 0$.

In the finite difference solution, when the $M + 1^{th}$ mesh point is just beginning to freeze, the most extreme temperature distribution that can arise is the steady state condition discussed above. This is:

$$\theta_m = mT/M \quad (m \leqslant M)$$
$$\theta_m = T \quad (m = M + 1)$$

Equation (2) written in terms of temperatures instead of heat content is

$$\theta_{m,r+1} = \theta_{m,r} + \gamma(\theta_{m+1,r} + \theta_{m-1,r} - 2\theta_{m,r}) \qquad (3)$$

where $\gamma = D\,.\,\delta t/(\delta x)^2$ and $D = k/c\rho$.

The boundary condition for θ is:

$$\left.\begin{aligned} \theta_{0,r} &= 0 \qquad (r \geqslant 0) \\ \theta_{M+1,r} &= T \qquad (r \geqslant 0) \\ \theta_{m,0} &= mT/M \qquad (m \leqslant M) \end{aligned}\right\} \qquad (4)$$

taking the time origin as the instant when the M^{th} mesh point became completely solid.

The steady state value to which the temperatures tend whilst the $M + 1^{th}$ mesh point is solidifying are:

$$\theta_m = mT/(M + 1) \quad (m \leqslant M + 1)$$

Putting $\theta_{m,r} = mT/(M + 1) + \phi_{m,r}$, equations (3) and (4) become

$$\left.\begin{aligned} \phi_{m,r+1} &= \phi_{m,r} + \gamma(\phi_{m+1,r} + \phi_{m-1,r} - 2\phi_{m,r}) \\ \phi_{0,r} &= 0 \\ \phi_{M+1,r} &= 0 \\ \phi_{m,0} &= mT/M(M + 1) \end{aligned}\right\} \qquad (5)$$

The solution of equations (5) is

$$\phi_{m,r} = \frac{T}{M(M+1)} \sum_{n=1}^{M} \frac{(-1)^{n-1}}{\tan n\pi/2(M+1)} \left[1 - 2\gamma\left(1 - \cos\frac{n\pi}{M+1}\right)\right]^r \sin\frac{n\pi m}{M+1} \qquad (6)$$

At the $M + 1^{th}$ mesh point, during freezing

$$H_{M+1,r+1} = H_{M+1,r} + k\,.\,\delta t(T + \theta_{M,r} - 2T)/\rho(\delta x)^2$$

where $\theta_{M,r} = MT/(M + 1) + \phi_{M,r}$

Put $\Delta H_{M+1,r} = H_{M+1,r+1} - H_{M+1,r}$

Therefore $$\Delta H_{M+1,r} = -\frac{k\,.\,\delta t}{\rho(\delta x)^2}\left(\frac{T}{M+1} - \phi_{M,r}\right) \qquad (7)$$

Let R_{M+1} be the number of δt required for solidification at the $M + 1^{th}$ mesh point, then

$$-\sum_{r=0}^{R_{M+1}-1} \Delta H_{M+1,r} = L \qquad (8)$$

Substituting equation (7) into equation (8)

$$\frac{k\,.\,\delta t}{\rho(\delta x)^2}\left(\frac{R_{M+1}T}{M+1} - \sum_{r=0}^{R_{M+1}-1} \phi_{M,r}\right) = L \qquad (9)$$

To complete the analytical solution of the finite difference scheme it is necessary to solve equations (6) and (9) for R_{M+1}. The present writers have found it necessary to introduce approximations in order to find R_{M+1}. The most advantageous approximation, in simplicity of analysis and accuracy of result, appears to be assuming that $\Sigma\phi_{M,r}$ is small compared with $R_{M+1}T/(M + 1)$. This leads to

$$R_{M+1} = L\rho(\delta x)^2(M + 1)/kT\,.\,\delta t$$
$$= L(M + 1)/cT\gamma \qquad (10)$$

So the total number of δt for the M^{th} mesh point to solidify is

$$\sum_{m=1}^{M} R_m = L \sum_{m=1}^{M} m/cT\gamma$$
$$= LM(M + 1)/2cT\gamma$$

The number of δt for the $M - 1^{th}$ mesh point to solidify is $LM(M - 1)/2cT\gamma$, so the mean time for the $M - 1^{th}$ and M^{th} mesh points to solidify is

$$\frac{LM^2\,.\,\delta t}{2cT\gamma} = \frac{L\rho x^2}{2kT}$$

This expression agrees with a simplified expression of Neumann's analytical solution as given by Carslaw and Jaeger.[4] The approximation $\Sigma\phi_{M,r} = 0$ is therefore sustained. Numerical evaluation shows that the difference between the time of evolution of half the latent heat, and half the time of evolution of all the latent heat, is very small.

Conditions governing the occurrence of the "steady state" error. Although accurate temperatures can be found from the solutions in which the steady state error occurs it may be more convenient to obtain a solution where such errors are

avoided. To this end the following criterion was developed. The solution of the differential form of equation (5) under the same boundary conditions is

$$\frac{2T}{M\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\exp\left[-\frac{Dn^2\pi^2t}{(M+1)^2(\delta x)^2}\right]\sin\frac{n\pi m}{M+1} \quad (11)$$

The Fourier expansion of $mT/(M+1)$ is

$$\frac{2T}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\sin\frac{n\pi m}{M+1}$$

So if the steady state does not arise when the $M+1^{th}$ mesh point is solidifying $\frac{1}{M}\exp\left[-\frac{n^2\pi^2L}{cT(M+1)}\right]$ must not be absolutely insignificant when compared with unity, for some value of n. The best test case is when $n=1$.

Then $$\exp\frac{\pi^2L}{cT(M+1)} \gg \frac{1}{M}$$

It is necessary to assign an arbitrary value to the expression "not absolutely insignificant." If this is taken as meaning a value greater than 1/20 the condition reduces to

$$\pi^2L/cT(M+1) < 3$$

or $$L/cT(M+1) < 0{\cdot}3$$

For an ingot cooling in a mould with a constant surface temperature of 1000° C, which may be expected,

$$L/(M+1)cT = 0{\cdot}4$$

when $M=1$, so the steady state is not expected to arise.

CONCLUSIONS

It is shown that the stability of numerical solutions of the heat flow equation is unaffected by the presence of a large isothermal latent heat evolution.

The finite difference method of solution results in a stepwise movement of the phase change front. Complete change of phase at any mesh point occurs when half the latent heat is evolved, as recorded by the numerical solution. This is approximately equal to half the calculated time for the evolution of all the latent heat.

An inaccuracy due to the setting in of the steady state in the conductor will not be significant if the dimensionless group

$$L/(M+1)cT \leqslant 0{\cdot}3$$

ACKNOWLEDGEMENTS

The authors wish to thank Prof. R. J. Sarjant for the advice and encouragement he has given in the course of this work.

REFERENCES

(1) PRICE, P. H., and SLACK, M. R. *Brit. J. Appl. Phys.*, **3**, p. 379 (1952).
(2) CRANK, J., and NICOLSON, P. *Proc. Cambridge Phil. Soc.*, **43**, p. 50 (1947).
(3) O'BRIEN, G. G., HYMAN, M. A., and KAPLAN, S. J. *Maths. Phys.*, **24**, p. 223 (1951).
(4) CARSLAW, H. S., and JAEGER, J. C. *Conduction of heat in solids* (London: Oxford University Press, 1947).

The β-ray absorption spectrum of $^{147}Pm_{61}$ and its application to thickness measurement

By L. MANDEL, B.Sc., Ph.D., A.Inst.P., Central Instrument Section, Imperial Chemical Industries Ltd., Welwyn

[*Paper received* 13 *November*, 1953]

The absorption in aluminium of the β-rays from $^{147}Pm_{61}$ has been measured and found to be nearly exponential, with an absorption length of 6·78 mg/cm^2 ± 3%. Curves are given which relate the measuring accuracy with absorber thickness in rapid response β-ray gauges using ^{147}Pm and ^{204}Tl as sources. These show that, under the same conditions, the use of ^{147}Pm leads to greater accuracy at thicknesses below about 20 mg/cm^2, and that, compared with ^{204}Tl, this source could usefully extend the lower limit of thickness measurement by a factor of about 4.

β-ray thickness gauges are now well established instruments, but the measurement of very thin materials—of, say, less than 10 mg/cm^2—with the β-ray sources so far available has nevertheless remained a problem. It has been shown[1] that the most commonly used isotope, namely ^{204}Tl, is not very suitable for measurements below about 10 mg/cm^2 and that the measuring accuracy falls rapidly below this thickness. Other fairly obvious choices, such as ^{45}Ca and ^{35}S, are either not available at sufficient source strength (i.e. of the order of 10 mc) or suffer from the disadvantage of a short life-time. An examination was therefore made of the characteristics of $^{147}Pm_{61}$ which is now becoming available as a fission product of uranium at a reasonable strength and in a safe form. It has a half-life of 2·6 years.

Although the β-ray energy spectrum of ^{147}Pm has been studied spectrometrically by a number of workers[2–4] who obtained almost linear Fermi plots, a direct measurement of the absorption spectrum was considered useful. The results show that the isotope is likely to reduce appreciably the lower limit of useful thickness measurement encountered with ^{204}Tl.

EXPERIMENTAL DETAILS AND RESULTS

The source for these measurements was supplied by the Radiochemical Centre, Amersham, and consisted of an evaporated solution of ^{147}Pm in hydrogen chloride, of approximately 1 mc strength, covering an area of about 1 cm^2. The β-rays were collected by an ionization chamber, using the arrangement shown in Fig. 1. The ion current flowing through the load resistance R ($10^{11}\ \Omega$) produced a potential drop, which was balanced out manually by means of the potentiometer P connected in series with R. The residual signal was fed into an electrometer amplifier E and thence to an indicating meter M. E contained three separate units, namely a vibrating reed electrometer, followed by an amplifier and a phase sensitive rectifier. A null reading of the meter M indicated that the potentiometer output equalled the potential drop developed across R.

The ionization chamber window consisted of approximately 6·5 mg/cm² aluminium foil, while the source was covered by 1 mg/cm² of mica. The air path from source to chamber amounted to about 4·5 mg/cm², which brought the total residual absorber to approximately 12 mg/cm². This might appear excessively high for measurements on ^{147}Pm, which

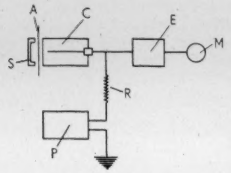


Fig. 1. The experimental scheme
S, source holder
A, absorbing material
C, ionization chamber
P, potentiometer
E, electrometer amplifier
M, indicating meter
R, load resistance

has a maximum β-ray energy of approximately 0·23 MeV[2-4], but it was thought helpful to approach the conditions to be found in an instrument, in order to assess the suitability of the source. The absorbing material *A* consisted of aluminium foil throughout.

The experimental points are shown in Fig. 2, where the fraction $\theta(x)$ of the maximum signal remaining in the presence of absorber of thickness x is plotted against x. For comparison, the absorption spectrum of ^{204}Tl was also measured

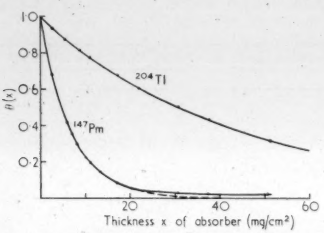


Fig. 2. A comparison of the absorption spectra of ^{147}Pm and ^{204}Tl
(Residual absorber approximately 12 mg/cm²)

under the same conditions and is shown on the same scale. It appears that the curve for ^{147}Pm, like that for ^{204}Tl, is very nearly exponential, with an absorption length of 6·78 mg/cm² aluminium ±3%, although it departs somewhat from the exponential form (shown by the broken line) for $x > 20$ mg/cm². This departure is quite likely due to a small contamination of ^{152}Eu and ^{154}Eu, which emit β-particles with energies up to approximately 1·9 MeV. It is possible that the spectra of later extractions of ^{147}Pm will not show this high energy tail.

THE ACCURACY OF THICKNESS MEASUREMENT USING ^{147}PM

We can now examine the range which might reasonably be covered by a thickness gauge using ^{147}Pm as source. In most β-ray instruments in which a reasonably rapid response is required, the random fluctuations of the output signal are the dominant factor limiting the accuracy of measurement. Under these conditions, it is convenient to define the limiting accuracy $\beta(x)$ as the ratio of the mean thickness x being measured to the smallest recognizable change of thickness, under the given conditions. It has already been shown[1] that $\beta(x)$ is very nearly related to $\theta(x)$ by the equation:

$$\beta(x) = -\sqrt{(2N_0\tau)}\frac{xd[\theta(x)]/dx}{\sqrt{\theta(x)}} \qquad (1)$$

provided the ratio of mean to mean square number of ion pairs formed in the chamber does not vary appreciably with x. N_0 is the average β-ray collection rate when $x = 0$ and τ is the effective integrating time constant of the electrical network.

Equation (1) has been applied to the data of Fig. 2, and the resulting curves relating the measuring accuracy with absorber thickness are shown in Fig. 3. The curve shown as a broken line, corresponds to that of Fig. 2, and is the form expected from a strictly exponential absorption spectrum.

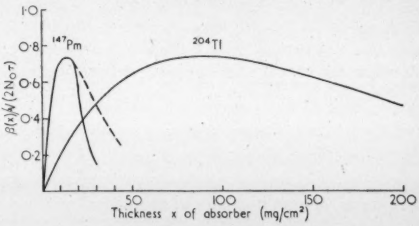


Fig. 3. A comparison of the measuring accuracy using ^{147}Pm and ^{204}Tl sources
(Residual absorber approximately 12 mg/cm²)

It will be seen that, with ^{147}Pm, the maximum accuracy is reached at less than one-sixth the absorber thickness corresponding to ^{204}Tl and that the $\beta(x)/\sqrt{(2N_0\tau)}$ curve rises approximately 6·5 times as steeply from zero. Now the 1 mc ^{147}Pm source yielded approximately one-third the ionization current obtained with 1 mc ^{204}Tl under the same conditions. When due allowance is made for this difference in the factor $\sqrt{(2N_0\tau)}$, it is not unreasonable to conclude that ^{147}Pm, available at millicurie strength, can usefully reduce the lower limit of β-ray thickness measurement by a factor $6{\cdot}5/\sqrt{3} \simeq 4$ compared with ^{204}Tl. We note further that, under similar conditions, the measuring accuracy $\beta(x)$ is greater with ^{147}Pm at thicknesses below about 20 mg/cm².

Apart from their uses as measuring instruments, β-ray gauges have industrial applications in the specific role of gauges for grading purposes, e.g. to test whether the dimensions of a product lie within some specification. It is well known that the output fluctuations of the instruments renders them liable to make errors in grading and these errors have been examined in some detail elsewhere.[5] It is sufficient to point out here that, since the effective signal fluctuations, under the same conditions of source strength and response time, are less with ^{147}Pm than with ^{204}Tl at thicknesses below about 20 mg/cm², the use of this source can bring about a considerable reduction of grading errors.

CONCLUSIONS

The β-ray absorption spectrum of ^{147}Pm in aluminium is found to be nearly exponential, with absorption length 6·78 mg/cm² ± 3%. The curves relating the measuring accuracy with thickness in reasonably fast response instruments using ^{147}Pm and ^{204}Tl sources show that, under the same conditions, ^{147}Pm leads to greater accuracy at thick-

nesses below about 20 mg/cm² and reaches its greatest accuracy at 13·5 mg/cm². Below 5 mg/cm², instruments using ^{147}Pm are four times more accurate than those using ^{204}Tl, under the same conditions.

ACKNOWLEDGEMENT

The writer wishes to thank Mr. Carter of the Radiochemical Centre, Amersham, for making an early sample of ^{147}Pm available for measurements.

REFERENCES

(1) MANDEL, L. *Brit. J. Appl. Phys.*, **5**, p. 58 (1954).
(2) LIDOFSKY, L., MACKLIN, P., and WU, C. S. *Phys. Rev.*, **76**, p. 1888 (1949).
(3) AGNEW, H. M. *Phys. Rev.*, **77**, p. 655 (1950).
(4) LANGER, L. M., MOTZ, J. W., and PRICE, J. H., JR. *Phys. Rev.*, **77**, p. 798 (1950).
(5) MANDEL, L. *J. Roy. Statist. Soc. B*, **16**. In the press (1954).

A method of investigating the secondary emissive properties of insulators

By C. N. W. LITTING, B.Sc., Ph.D., A.M.I.E.E., A.Inst.P., Electrical Engineering Laboratories, University of Manchester

[*Paper first received* 21 *October*, 1953, *and in final form* 2 *April*, 1954]

Methods are described for measuring the secondary emission ratio (δ) of an insulator together with the energy distribution of the emitted electrons. The secondary emission ratio is obtained by means of a type of electron beam voltmeter which is used to measure the potential of the bombarded surface thus allowing the rate of change of charge to be determined. The results show good agreement with those obtained by another method and also check with d.c. measurements on metal surfaces. The energy distribution of the secondary electrons from an insulator surface is measured using the electron beam voltmeter and also by a method employing cathode potential stabilization.

As the electron beam voltmeter method for measuring δ depends on the ratio of two time measurements it should be capable of greater accuracy than methods which involve the measurement of output pulses from an amplifier.

Most previous measurements of the secondary emission ratio (δ) of insulators have been made by methods in which the insulator has been treated as a poor conductor. Using these methods the actual potential of the insulator surface is somewhat doubtful. Some previous investigations have also been made which do not require conduction in the insulator but use dynamic methods.[1–6] All these methods are limited to the region where $\delta > 1$, and depend on the measurement of the output pulse from a high-gain amplifier. At low levels in the presence of noise these methods are necessarily inaccurate. In addition, some of the methods[4,5,6] employ a modulated beam and therefore any "cloud pulses" due to the cloud of space charge produced in the vicinity of the screen would interfere with the results.

The method to be described is also limited to the region where $\delta > 1$. It depends on the ratio of two time measurements and should thus be capable of greater accuracy, especially in the region where δ is not much greater than unity.

A dynamic method will also be described for determining the energy distribution of the secondary electrons. Two previous workers have employed dynamic methods to obtain this distribution but their results are affected by any cloud pulses due to space charge effects[5,6] and in addition they depend on the shape of the output wave-form from a high-gain amplifier. Results obtained by static methods are suspect because of the uncertainty of the actual potential of the surface.

MEASUREMENT OF SECONDARY EMISSION RATIO

Description of the method. The method employs a type of electron beam voltmeter to determine the potential of a surface.[7,8] The system is illustrated in Fig. 1. The vertical electron beam in the diagram—called the "write" beam—is rectangular in cross-section and covers the whole of the insulator surface; this beam causes the surface of the insulator to stabilize at approximately mesh potential provided that $\delta > 1$. The horizontal beam—the "read" beam—which passes close to the surface and then impinges on a fluorescent screen, is used to determine the potential difference between the surface and the mesh.

A square wave of 100 V amplitude is applied to a backing plate behind the insulator surface, thus altering the potential of the surface and the electric field between the surface and the mesh. If the surface is positive with respect to the mesh

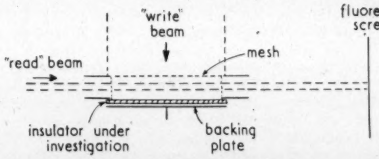


Fig. 1. Electron beam voltmeter

it will gain electrons from the "write" beam and charge towards the mesh at a rate proportional to the "write" beam current (I_B). Alternatively, if the mesh is positive, the surface will suffer a net loss of electrons (provided $\delta > 1$) and charge positive at a rate proportional to $(\delta - 1)I_B$.

If now the mark to space ratio of the square wave is altered, the relative loss and gain of charge during the positive and negative excursions of the backing plate is also altered. Thus provided $(\delta - 1)$ is positive it can be arranged that there is no net loss of charge from the surface over a complete cycle, and that the potential excursions of the surface above and below its original stable potential are equal. This is illustrated in Fig. 2. Then if t_1 is the time during which the backing plate is positive and t_2 the time during which it is negative, and as the charge gained by the surface whence it is positive equals that lost when it is negative, it follows that

$$I_B t_1 = (\delta - 1) I_B t_2$$

or

$$\delta = (t_1 + t_2)/t_2$$

Thus, in order to measure the secondary emission ratio, it is merely necessary to vary the mark to space ratio of the wave-form applied to the backing plate until the display on the fluorescent screen goes equal distances above and below the line obtained when the backing plate is at a steady potential. Then, from the timing of the wave-form, δ is easily obtained.

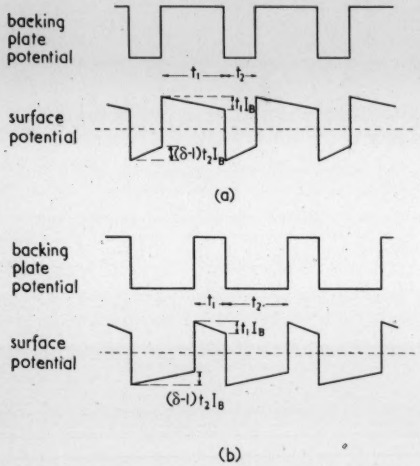


Fig. 2. Wave-forms in a "balanced" state obtained using the electron beam voltmeter method
(*a*) $\delta > 2$, (*b*) $\delta < 2$.

One disadvantage of this method is the presence of the mesh in the path of the primary beam. When the mesh is positive with respect to the surface under test, none of the secondaries from the mesh will reach the surface, and therefore they will not interfere with the charging process already described. However, when the mesh is negative, secondaries from the mesh will be attracted to the surface and the rate of negative charging of the surface will be increased. The presence of the mesh will therefore result in the measured value of δ being too low.

In order to prevent the secondaries from the mesh reaching the surface, the mesh was replaced by a double-grid structure. The first grid consisted of parallel wires immediately behind which was placed an alined grid consisting of wires of half the diameter. If the second grid, which acts as the deflector plate for the "read" beam, is held 100 V negative with respect to the first grid, no secondaries from the first grid should be able to pass it and reach the surface under test. With the grid spacings used it can be shown that, unless the voltage between the two grids exceeds four-fifths of the accelerating voltage between the cathode and the first grid, none of the electrons in the primary "write" beam will be deflected sufficiently on passing through the first grid to strike the second grid.[9] Thus only the primary beam electrons should be able to reach the surface and true values of δ can be obtained.

As an alternative to pulsing the backing plate, the mesh (or alined grids) could be pulsed, the backing plate being earthed. If, however, the potential of the mesh is altered, there is no guarantee that the beam current passing through it will not be changed. This is especially true when using low values of primary electron velocity. For this reason this method was not used.

Measurements were also taken by a method similar to the commutated collector method employed by Schagen.[3] In this case a 50 V pulse was applied to the mesh, the time period being sufficiently long (500 to 5000 μsec) to allow the surface to attain equilibrium in each state. After each change in mesh potential the surface charges towards its new equilibrium potential, producing a current pulse in the lead from the backing plate of amplitude initially proportional to either $(\delta - 1)I_B$ or I_B depending on whether the mesh has been pulsed positively or negatively. The current output from the backing plate is shown in Fig. 3, as is also

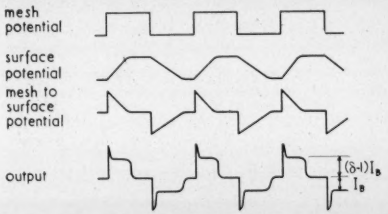


Fig. 3. Wave-forms obtained using the commutated collector method

the change in potential of the surface, together with the potential difference between the mesh and the surface. This latter potential difference is displayed on the fluorescent screen by means of the "read" beam. In actual fact, the pulses due to direct pick-up are clipped by germanium diodes to prevent overloading the current amplifier. The type of output actually obtained from the backing plate is shown in Fig. 4. The ratio of the amplitudes of the two steps gives the value of $(\delta - 1)$.

Fig. 4. Determination of secondary emission ratio using the commutated collector method

It should be noted that a 50 V pulse was used in this latter method. This is because the period of the pulses is sufficiently long to allow the surface to reach equilibrium, and, therefore, immediately after each edge of the mesh wave-form there is a potential difference of ± 50 V between the mesh and the surface. In the electron beam voltmeter method a 100 V pulse is used, for in that case the pulse lengths are much shorter ($t_1 + t_2 = 46$ μsec) so that the

surface does not charge very far during a positive or negative period and in the final "balanced" state there is again a difference of ± 50 V between the surface and the mesh. Thus in both cases, the measurements are being made under similar conditions. The measurements made by the commutated collector method will be inaccurate at low accelerating potentials, since the beam current will vary slightly with the potential applied to the mesh.

Experimental results. The secondary emission coefficient is plotted as a function of the accelerating voltage for surfaces of magnesium oxide, willemite, mica and Pyrex glass in Fig. 5. The magnesium oxide surface was prepared by

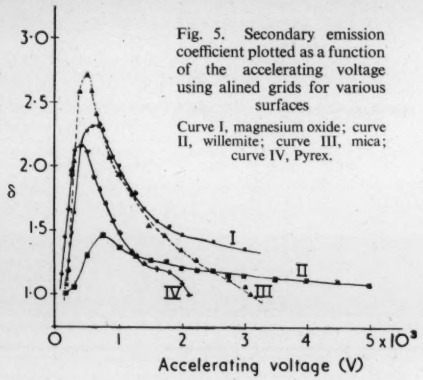


Fig. 5. Secondary emission coefficient plotted as a function of the accelerating voltage using alined grids for various surfaces

Curve I, magnesium oxide; curve II, willemite; curve III, mica; curve IV, Pyrex.

holding a glass plate in the smoke from a burning piece of magnesium, and the willemite surface was prepared by a pouring technique using a non-permanent nitro-cellulose binder. These curves were obtained by the electron beam voltmeter method using the alined grids; results obtained with a normal mesh showed much lower values of δ, as would be expected. Experiments with the double-grid structure indicated that under the worst conditions the second grid captured only 1% of the current passing through it. The

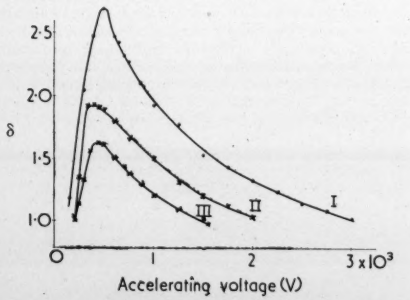


Fig. 6. Secondary emission coefficient plotted as a function of the accelerating voltage using alined grids for a sample of mica

Curve I, immediately after insertion; curve II, 2 h after insertion; curve III, 4 days after insertion.

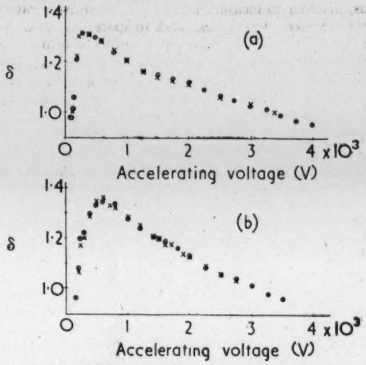


Fig. 7. Secondary emission coefficient plotted as a function of the accelerating voltage using normal mesh grids for (*a*) evaporated magnesium and (*b*) evaporated copper

○ = d.c. method; × electron beam voltmeter method.

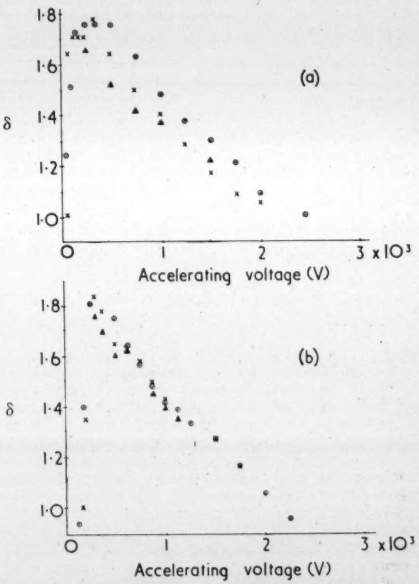


Fig. 8. Secondary emission coefficient plotted as a function of the accelerating voltage using alined grids for (*a*) evaporated copper and (*b*) nickel

○ = d.c. method; × = electron beam voltmeter method; ▲ = commutated collector method.

electrons which reached the second grid appeared to be high velocity secondaries (>100 V) released from the first grid and not actually electrons in the primary beam.

The measurements were performed using a continuously-pumped demountable system and therefore repeatable results could not be obtained from apparently similar surfaces; in fact the value of δ for any particular surface was found to vary with time. This is illustrated in Fig. 6 which shows results obtained with a sample of mica over a period of four days.

In order to check the method, measurements were taken on evaporated metal surfaces deposited on an insulator. No electrical contact was made to the surface and pulses were applied to a backing plate, measurements being taken as with an insulator. A d.c. connexion was also made to the surface so that δ could be measured either by the electron beam voltmeter method or by a normal d.c. method. Agreement between the two methods was extremely good as illustrated in Fig. 7 which shows results obtained using a simple mesh.

In addition, measurements could be taken using the commutated collector method. Results obtained by the three methods using the alined grids structure are shown in Fig. 8. When comparing results obtained by different methods it is, of course, necessary that the results should be taken as closely together in time as possible.

The effect of the angle of incidence of the primaries could be demonstrated by rotating the "read" gun, as the mesh and insulator surface were both mounted on that gun. Results for a magnesium oxide surface using a simple mesh are shown in Fig. 9.

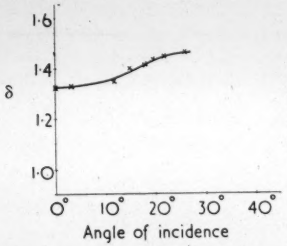


Fig. 9. Effect of angle of incidence on secondary emission coefficient using a normal mesh grid for a magnesium oxide surface $V_p = 2000$ V.

MEASUREMENT OF THE ENERGY DISTRIBUTION OF THE SECONDARY ELECTRONS

Two methods were used to investigate the energy distribution of the secondary electrons from the surface of an insulator. One method depended on cathode potential stabilization of the surface whilst the other used the electron beam voltmeter technique. Unfortunately, the two methods could not be used in the same tube on the same surface and therefore it has not been possible to use one method to check the results obtained by the other. As the two methods involve different electrode structures they will be described separately.

Method employing cathode potential stabilization. The potential of the surface is stabilized at the cathode potential of an electron gun which "illuminates" the whole of the surface with a high-current density electron beam (Fig. 10). A second gun produces an interrupted low-current density beam which is focused on the surface. As the current density is low it will not appreciably effect the stabilization of the surface by the low-velocity beam. A Faraday cage is arranged so that it can receive secondaries from the surface, but not those produced at any of the electrodes of the second electron gun. The Faraday cage is a.c.-coupled to a high-gain current amplifier. Thus the amplitude of the pulses of current due to secondary electrons produced by the interrupted beam may be measured as a function of the direct potential of the cage. The percentage of the secondaries with energies above any given value may therefore be obtained.

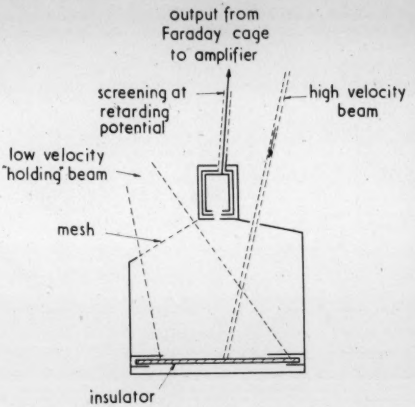


Fig. 10. Electrode system used to obtain the energy distribution of secondary electrons by the cathode potential stabilization method

Unfortunately, although cathode potential stabilization is sufficiently accurate for the measurement of secondary emission ratios,[2] the potential of the surface is not defined as accurately as could be desired for the present investigation. The actual potential difference between the surface and the Faraday cage may differ by a few volts from that measured between the surface and the cathode of the "holding" gun. Thus the method is not very successful as it is difficult to determine the allowance to be made for this effect. Typical curves obtained for willemite are shown in Fig. 11. In these curves the percentage of the electrons reaching the Faraday cage is plotted as a function of the retarding voltage actually measured (i.e. not corrected). The potential at which it is

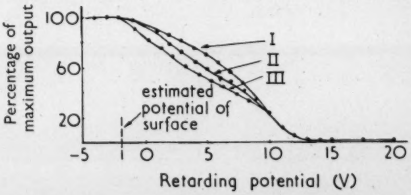


Fig. 11. Energy distribution of secondary electrons from willemite using the cathode potential stabilization method

Curve I, $V_p = 1000$ V; curve II, $V_p = 1500$ V; curve III, $V_p = 2000$ V.

estimated that the surface is stabilized is indicated on the curves. It is probable that the current to the Faraday cage at retarding potentials greater than 20 V is due to secondary electrons having an emission energy equal to that of the primary electrons striking the surface.

Results obtained for nickel, treating it as an insulator, are shown in Fig. 12. In these curves the corrected value of retarding potential is used. Results obtained by a standard d.c. retarding potential method for the same nickel surface are also shown in Fig. 12 and it will be seen that they agree with those obtained when treating the nickel as an insulator.

Although the method appears to give correct results for a nickel surface treated as an insulator, it does not necessarily follow that the same applies to an insulator surface. It is doubtful whether the cathode potential stabilization actually functions at the spot bombarded by the second electron beam and therefore the actual potential of that point is likely to be different from that of the surrounding surface. This, of course, would not apply in the case of nickel as the whole

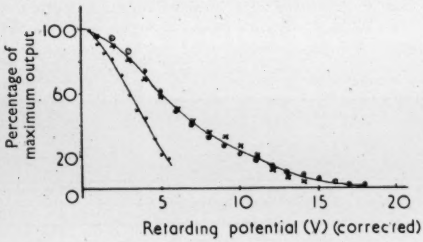


Fig. 12. Energy distribution of secondary electrons from nickel

○ = d.c. method, $V_p = 1000$ V; × = cathode potential stabilization method, $V_p = 1000$ V; • = electron beam voltmeter method (alined grids), $V_p = 800$ V.

surface would then be at the same potential. Thus even if the potential of the surrounding surface was known accurately —which it is not—the true retarding potential would be unknown and would have to be estimated. On the other hand, if the stabilization was effective so that no local potential changes occurred on the surface, some of the electrons from the holding beam which might otherwise have reached the collector will be required to maintain the potential of the bombarded spot. This would give rise to an output of opposite sign superposed on the desired output but, as the velocity of the electrons in the holding beam is so low, the effect would only cause an error at very low values of retarding potential.

Taken on the whole, the results obtained by this method are, therefore, of a somewhat doubtful value. However, as previously mentioned, another method has been developed and will now be described.

Method employing the electron beam voltmeter. The "write" beam which covers the whole of the surface is pulsed on and off and the secondaries are collected by a Faraday cage to which d.c. retarding potentials are applied. The output from the cage is a.c.-coupled to an amplifier so that the amplitude of the current pulses may be measured. The "read" beam is used merely to determine the potential of the surface in order to obtain the true retarding potential.

This method has not been thoroughly investigated but results have been taken with the electrode structure illustrated in Fig. 13. It will be noted that the collector is not a Faraday cage but merely a metal plate. Results obtained for nickel using this preliminary structure are shown in Fig. 12. Due to the background noise it was impossible to take readings with higher retarding potentials. It did, however, appear

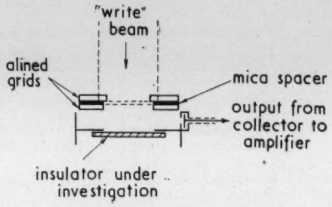


Fig. 13. Electrode system used to obtain the energy distribution of secondary electrons by the electron beam voltmeter method

that at higher values of the retarding potential the current to the negative collector reversed in sign. This would appear to be due to tertiary electrons which are liberated at the collector by high-energy secondaries and escape to the surrounding metal. In order to prevent this effect it will therefore be necessary to use a Faraday cage as collector in all future experiments.

A disadvantage of this method is that the secondaries measured are not those emitted approximately normal to the surface, as in the previous method. Thus as the electrons are not emitted in a field free space, the number of electrons of any given energy arriving at the collector will be, to some extent, a function of the angular distribution of the secondaries.

The method has the advantage that the whole of the insulator surface is at the same potential and that this surface potential is measured directly.

CONCLUSIONS

From the various results it can be seen that the electron beam voltmeter method is suitable for measuring the secondary emission coefficient of insulators for various angles of incidence in the region where ($\delta > 1$). It can also be used to determine the energy distribution of the emitted secondaries, although there will be some error at low energies as the secondaries are not emitted into a field free space.

In order to obtain consistent results it is necessary to work under better vacuum conditions than can be obtained in a demountable vacuum system.

REFERENCES

(1) Heimann, W., and Geyer, K. *Elektrische Nachrichten-Technik*, **17**, p. 1 (1940).
(2) Salow, H. *Z. Tekhn Phys.*, **21**, p. 8 (1940).
(3) Schagen, P. *Philips Res. Rep.*, **6**, p. 135 (1951).
(4) Heydt, H. L. *Rev. Sci. Instrum.*, **21**, p. 639 (1950).
(5) McKay, K. G. *J. Appl. Phys.*, **22**, p. 89 (1951).
(6) Barthélemy, R. *C.R. Acad. Sci.* [*Paris*], **232**, p. 20 (1951).
(7) Piore, E. R., and Morton, G. A. *J. Appl. Phys.*, **11**, p. 153 (1940).
(8) Williams, F. C., Kilburn, T., Litting, C. N. W., Edwards, D. B. G., and Hoffman, G. R. *Proc. Instn Elect. Engrs*, II, **100**, p. 523 (1953).
(9) Jonker, J. L. H. *Philips Res. Rep.*, **6**, p. 372 (1951).

Requirements contributing to the design of devices used in correcting electron lenses

By G. D. ARCHARD, A.Inst.P., Associated Electrical Industries, Ltd., Aldermaston, Berks.

[*Paper received* 13 *January*, 1954]

The voltage requirements for a modification of the spherical correction system of Seeliger[2] are given. The potential changes introduced by unintentional asymmetries of non-rotationally-symmetrical elements are calculated and an expression is given for the resulting aberrations. It is shown that some of these can be removed by the Stigmator of Rang,[4] but that others remain which would necessitate positioning to 1 μ in order to attain a resolution improvement of three times. With the addition of certain other elements, tolerances could be raised to 10 μ for a resolution improvement of ten times. In the course of the work, the mechanism of Seelinger's system and its suggested modification are compared with another system proposed by Burfoot.[5]

It was shown theoretically by Scherzer[1] that the chromatic aberration of electron lenses could be corrected by means of non-rotationally-symmetric* elements producing a potential of the form $\Phi_2(z)r^2\cos 2\theta$ [Fig. 1(*a*)], and that spherical aberration could be corrected by a suitable combination of these with other elements producing a potential of the form $\Phi_4(z)r^4\cos 4\theta$ [Fig. 1(*b*)]. Seeliger[2] constructed such a spherical corrector, but pointed out that marked improvement in resolution might not be attained without increased mechanical and electrical stability. The limitation of resolution by mechanical inaccuracies in (magnetic) *round* lens construction has been theoretically investigated by Sturrock,[3] who showed that with normal tolerances (variations of 1 μ in diameter of bore) there would be a first order astigmatism which would completely mask spherical aberration; in practice, such astigmatism may largely be removed by means of the Stigmator of Rang,[4] which basically comprises a "Φ_2" unit capable of rotation about the axis. (A Φ_2 unit by itself converges rays in one plane and diverges those in the plane at right angles; together with a round lens it therefore produces a controllable difference in focal length in the two planes.) Burfoot[5,6] has given the theory of an electrostatic lens made spherically correct *internally* by so shaping the electrodes as to incorporate Φ_2 and Φ_4 potential components, but the calculated mechanical tolerances would be extremely difficult to fulfil. It is therefore desirable to investigate fundamentally the voltage requirements and positioning tolerances of *external* correctors such as those of Seeliger, while at the same time considering whether there is any more convenient way in which they can be arranged.

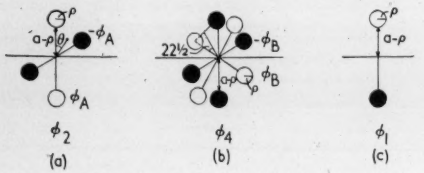


Fig. 1. Simple forms of non-rotationally-symmetrical elements used in correction of spherical aberration

1. VOLTAGE REQUIREMENTS

Comparison of spherical correction systems

Seeliger's system is illustrated in Fig. 2. On leaving the lens to be corrected, the electron beam is regarded as travelling parallel to the axis. A cylindrical lens focuses the beam in the *y*-direction and a correcting "Φ_4" element is inserted at the line focus. A round lens then makes the *y*-component of the rays parallel to the axis and focuses the beam in the *x*-direction. Another correcting element is placed at the resulting line focus. A second cylindrical lens makes the *x*-components of the rays parallel without affecting the *y*-components, and a third correcting element is placed in the parallel beam. Each cylindrical lens combines the functions of a *convergent* round lens and a "Φ_2" unit (Fig. 3): this concept will be useful for comparison with other systems.

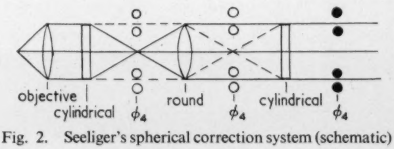


Fig. 2. Seeliger's spherical correction system (schematic)
R = round lens.
– – – – *xz* plane; ——— *yz* plane.

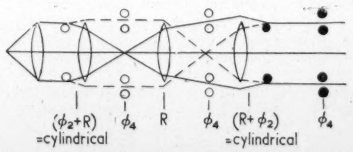


Fig. 3. Modification of Fig. 2, eliminating cylindrical lenses
R = round lens.
– – – – *xz* plane; ——— *yz* plane.

Burfoot's arrangement is illustrated schematically in Fig. 4. The first and third of the shaped electrodes may each be regarded as a combination of a *divergent* (aperture) lens and a Φ_2 unit, while the second electrode may be regarded as a round lens superimposed on a pair of Φ_4 units. These correspond respectively to Seeliger's two cylindrical lenses, and his central round lens plus two Φ_4 units. Burfoot's fourth electrode corresponds to Seeliger's third Φ_4 unit. It is Burfoot's effective use of *divergent* round components which makes it unnecessary for his rays to cross the axis forming real line foci.

Seeliger's system has been drawn for rays emerging parallel to the axis, that is, for infinite magnification. If the fact of

* "Rotationally-symmetric" will hereinafter be written "round."

finite magnification is taken into consideration, the system may be simplified (Fig. 5). Hypothetical rays, x_α, y_α (starting from the axial point of the object at inclinations $x'_\alpha = y'_\alpha = 1$) leave the lens under correction at inclinations $x'_\alpha = y'_\alpha =$

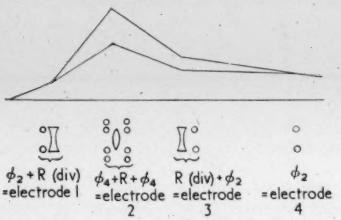


Fig. 4. Burfoot's spherical correction system (schematic)
R = round lens.

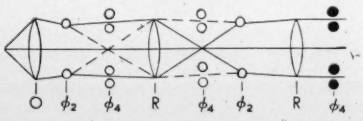


Fig. 5. Modification of Fig. 3 for non-infinite magnification
R = round lens.
- - - - x_α (xz plane); ——— y_α (yz plane).

$1/M$ (M = magnification). A Φ_2 unit then separates them; conveniently, x'_α is doubled and y'_α made zero. This means, in effect, that the round lens part of Seeliger's first cylindrical lens has been absorbed into the lens under correction. If the latter is the objective of an electron microscope, the round lens part of Seeliger's second cylindrical lens may be regarded as an intermediate projection stage contributing a little magnification (it cannot be absorbed into the projector proper as its focal length cannot be made short enough).

For the purpose of ensuing numerical calculations, it will be assumed that the modification suggested has been carried out, so that regions of finite Φ_2 and finite Φ' do not interpenetrate (Φ is the axial potential due to the objective and the other round lenses, and dashes indicate differentiation with respect to z).

Application of spherical correction conditions

Consider Scherzer's equations for the removal of spherical aberration [equation 4.5 in Ref. (1)]. In the present case (Figs. 3 or 5) the whole arrangement is symmetrical and Φ_2, Φ' do not interpenetrate, so that the conditions reduce to:

$$C_s + \int_{z_0}^{z_i} \left(\frac{5}{6}\frac{\Phi_2^2}{\Phi^2} - 2\frac{\Phi_4}{\Phi}\right) x_\alpha^4 dz = 0 \qquad (1)$$

$$C_s + \int_{z_0}^{z_i} \left(-\frac{1}{2}\frac{\Phi_2^2}{\Phi^2} + 6\frac{\Phi_4}{\Phi}\right) y_\alpha^2 x_\alpha^2 dz = 0 \qquad (2)$$

$$C_s + \int_{z_0}^{z_i} \left(\frac{5}{6}\frac{\Phi_2^2}{\Phi^2} - 2\frac{\Phi_4}{\Phi}\right) y_\alpha^4 dz = 0 \qquad (3)$$

where C_s comprises the spherical aberration constant of the objective plus small contributions from the other round lenses.

If the Φ_2 and Φ_4 units may be regarded as thin (Fig. 3), x_α and y_α may be assigned constant values over any particular unit; thus, over the Φ_2 units x_α and y_α may be regarded as equal and of the order of the objective focal length f, while at the first and second Φ_4 units one of them may be regarded as zero and the other equal to f.

These equations show why it would be impossible for correction to occur if the Φ_4 units were placed where $x_\alpha^2 = y_\alpha^2$, for then equations (1) and (2) would yield:

$$4C_s + \int_{z_0}^{z_i} 2\frac{\Phi_2^2}{\Phi^2} x_\alpha^4 dz = 0 \qquad (4)$$

which could not be fulfilled as C_s is always positive.

But if $x_\alpha = y_\alpha$ over the Φ_2 units, as assumed, equations (1) and (3) give:

$$\int_{z_0}^{z_i} \frac{\Phi_4}{\Phi} x_\alpha^4 dz = \int_{z_0}^{z_i} \frac{\Phi_4}{\Phi} y_\alpha^4 dz \qquad (5)$$

This condition may be satisfied if the Φ_4 units at the zeros of x_α, y_α respectively are made equal. (It may be noted that these positions for Φ_4 are as far as possible removed from the forbidden condition of $x_\alpha^2 = y_\alpha^2$.)

Condition (2) could now be satisfied without the addition of the third Φ_4 unit simply by making:

$$\frac{1}{2}\int_{z_0}^{z_i} \frac{\Phi_2^2}{\Phi^2} x_\alpha^2 y_\alpha^2 dz = C_s \qquad (6)$$

but insertion of numerical values shows that the potentials required for the Φ_2 units would be of the order of the accelerating potential, so that in general the third Φ_4 unit *will* be required.

Form of Φ_n units

The simplest geometrical arrangement leading to potentials of the form Φ_2, Φ_4 comprises symmetrical arrays of spheres grouped round the axis as in Fig. 1. Suppose alternate spheres in Fig. 1(*a*) support charges $\pm q$. Potential at (r, θ, z) is:

$$\begin{aligned} &q/\sqrt{[(a - r\cos\theta)^2 + r^2\sin^2\theta + z^2]} \\ &\quad + q/\sqrt{[(a + r\cos\theta)^2 + r^2\sin^2\theta + z^2]} \\ &\quad - q/\sqrt{[(a - r\sin\theta)^2 + r^2\cos^2\theta + z^2]} \\ &\quad - q/\sqrt{[(a + r\sin\theta)^2 + r^2\cos^2\theta + z^2]} \end{aligned} \qquad (7)$$

Provided the electrons do not approach the spheres (i.e. $r \ll a$), these terms may be expanded binomially and combined to form:

$$\frac{3qa^2r^2\cos 2\theta}{(a^2 + z^2)^{5/2}} \qquad (8)$$

If the potentials on the spheres are $\pm\Phi_s$ and their radii are ρ:

$$q \simeq \Phi_s\rho \qquad (9)$$

so that in Scherzer's notation:

$$\Phi_2 = 3\Phi_s\rho a^2/(a^2 + z^2)^{5/2} \qquad (10)$$

Similarly, it may be shown that the eight spheres of Fig. 1(*b*) (alternately $\pm\Phi_s$) lead to a distribution $\Phi_4 r^4 \cos 4\theta$ where:

$$\Phi_4 = 35\Phi_s\rho a^4/8(a^2+z^2)^{9/2} \qquad (11)$$

At a later stage it will be necessary to consider "Φ_1" units [Fig. 1(*c*)]. Similar considerations lead to a potential distribution of:

$$\frac{\Phi_s\rho}{(a^2+z^2)^{1/2}}\left[\frac{2ax}{(a^2+z^2)} - \frac{3ax(x^2+y^2)}{(a^2+z^2)^2} + \frac{5a^3x^3}{(a^2+z^2)^3}\right] \qquad (12)$$

where $x = r\cos\theta$, $y = r\sin\theta$.

Simple arrays of spheres such as these would be inconvenient to construct as well as inefficient in their action. It will be shown in a later paper that the same value of Φ_2, etc., can be obtained with much smaller values of Φ_s by expanding the spheres into rods parallel to the axis [Fig. 6(*a*)], or to a

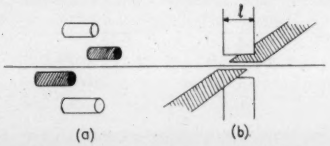


Fig. 6. (*a*) Rod form of Φ_2 unit. (*b*) Laminar form of Φ_2 unit

lesser extent by expanding them into rods perpendicular to the axis. Seeliger used for his Φ_4 units an array of eight thin laminae [of which the Φ_2 analogue is shown in Fig. 6(*b*)]; these are again more efficient than spheres. Nevertheless, it will be convenient to treat arrays of spheres first and afterwards to consider the improvements which may be made by altering the geometry.

Potentials required for spherical aberration correcting systems

Suppose that the values of Φ_s for the various units (Figs. 3 or 5) are:

For the pair of Φ_2 units: $\Phi_{s(2)}$
For the pair of Φ_4 units: $\Phi_{s(4p)}$ (p for pair)
For the third Φ_4 unit: $\Phi_{s(4t)}$ (t for third)

Taking $x_\alpha \simeq y_\alpha \simeq f$ over the two Φ_2 units and the third Φ_4 unit, $x_\alpha \simeq f$, $y_\alpha \simeq 0$ over the first Φ_4 unit, and $x_\alpha \simeq 0$, $y_\alpha \simeq f$ over the second Φ_4 unit, equations (1) and (2) yield:

$$C_s + \frac{5}{3}\frac{\Phi^2_{s(2)}}{\Phi^2_A}f^4\int\left[\frac{3a^2\rho}{(z^2+a^2)^{5/2}}\right]^2 dz - 2\frac{\Phi_{s(4p)}+\Phi_{s(4t)}}{\Phi_A}f^4\int\frac{35a^4\rho}{8(z^2+a^2)^{9/2}}dz = 0 \qquad (13)$$

$$C_s - \frac{\Phi^2_{s(2)}}{\Phi^2_A}f^4\int\left[\frac{3a^2\rho}{(z^2+a^2)^{5/2}}\right]^2 dz + 6\frac{\Phi_{s(4t)}}{\Phi_A}f^4\int\frac{35a^4\rho}{8(z^2+a^2)^{9/2}}dz = 0 \qquad (14)$$

where Φ_A is the accelerating potential.

Integration is taken in each case "over the unit" in question, which implies, as has already been assumed, that the units are thin, i.e. that their effect is chiefly concentrated within a distance a of their centres. Hence, for approximate purposes, little difference is made by setting the limits of integration $\pm\infty$, and by a slight transformation deducing:

$$C_s - 7{\cdot}73\frac{\Phi^2_{s(2)}}{\Phi^2_A}\frac{f^4\rho^2}{a^5} + 24{\cdot}0\frac{\Phi_{s(4t)}}{\Phi_A}\frac{f^4\rho}{a^4} = 0 \qquad (15)$$

$$4C_s + 31{\cdot}0\frac{\Phi^2_{s(2)}}{\Phi^2_A}\frac{f^4\rho^2}{a^5} - 24{\cdot}0\frac{\Phi_{s(4p)}}{\Phi_A}\frac{f^4\rho}{a^4} = 0 \qquad (16)$$

Φ_4 *potentials.* Now in principle $\Phi_{s(2)}$ may be made indefinitely weak simply by making the apparatus indefinitely long (Fig. 5), and in practice it will be found that the $\Phi_{s(2)}$ terms in equations (15) and (16) may be neglected. Then approximately:

$$\Phi_{s(4t)}/\Phi_A = -\Phi_{s(4p)}/4\Phi_A = -a^4C_s/24f^4\rho \qquad (17)$$

Typical values might be: $a = 0{\cdot}4$ cm, $\rho = 0{\cdot}1$ cm [Fig. 1(*b*) shows that ρ must be less than $a \sin 22\frac{1}{2}°$], $f = 0{\cdot}5$ cm $= C_s$, $\Phi_A = 50$ kV. These would give:

$$\Phi_{s(4t)} = -4\text{ kV}$$
$$\Phi_{s(4p)} = 16\text{ kV}$$

and as, with the dimensions given, adjacent spheres would be separated by about 1 mm, insulation would present difficulty.

It will be shown in the later paper that the use of a group of laminae (such as those of Seeliger) would, with the same value of a, produce an improvement factor of the order of:

$$l/2\pi\rho \qquad (18)$$

where l equals depth of lamina (parallel to the axis); thus, laminae with $l = 2$ cm would require potentials some three times less than those given above. Considerations of space might make this value of l inconveniently large, but another advantage lies in the fact that the laminae can be made very thin (say 0·01 cm) so that a may be reduced. A reduction of a by a factor of 2 would further reduce the potentials required by a factor of 16, while at the same time spacing between adjacent electrodes would be *increased* to 1·5 mm. Hence, there should be no difficulty in making an overall reduction of the above potential values by a factor of 20. The compensating disadvantage is that tolerances are also dependent on a, as will appear later.

Φ_2 *potentials.* It follows immediately from Scherzer's ray equations that for a pure Φ_2 unit:

$$x''_\alpha + (\Phi_2/\Phi_A)x_\alpha = 0 \qquad (19)$$

$$y''_\alpha - (\Phi_2/\Phi_A)y_\alpha = 0 \qquad (20)$$

and for a thin unit x_α and y_α may be replaced by f. Hence, change in slope x'_α produced by the Φ_2 unit is:

$$\pm(f/\Phi_A)\int\Phi_2 dz \qquad (21)$$

In order to produce the ray system of Fig. 5, the first Φ_2 unit must change the slope by $1/M$, where M is the magnification of the lens under correction. Hence:

$$\Phi_{s(2)}/\Phi_A = a^2/4\rho fM \qquad (22)$$

With the previous values of a, f, ρ and $M = 20$ this yields:

$$\Phi_{s(2)} = 2\text{ kV}$$

with a separation of 3·6 mm between adjacent spheres. As in the case of Φ_4 units, this potential would be considerably reduced by the use of laminae.

The central round lens

This has not been mentioned in detail because the construction and properties of round lenses are well known. It is clear from Fig. 5, however, that its focal length will have to be of the order of one quarter of the distance between the Φ_2 units.

2. POSITIONING TOLERANCES

In view of the severe tolerance limitations of Burfoot's system of spherical correction, it is desirable to investigate fundamentally the tolerances of units of the Φ_2 and Φ_4 form considered in Part 1 of the Conclusions.

Expressions (10) and (11) have already been given for the potential at (r, θ, z) for such groups of spheres, and now the effect of displacing individual spheres from their position of symmetry will be considered.

Effect of displacing one sphere radially

In either Fig. 1(*a*) or (*b*) let one of the spheres in the xz plane be displaced radially δa. Difference of potential thus introduced at (r, θ, z) will be:

$$\frac{q}{(a^2+z^2)^{\frac{1}{2}}}\left\{\left[1+\frac{r^2-2(a+\delta a)r\cos\theta+2a\delta a}{(a^2+z^2)}\right]^{-\frac{1}{2}}-\left[1+\frac{r^2-2ar\cos\theta}{(a^2+z^2)}\right]^{-\frac{1}{2}}\right\} \quad (23)$$

If consideration is restricted to aberrations of third order, it is sufficient to extract powers of r up to the fourth. Writing then $x = r\cos\theta, y = r\sin\theta$, expression (23) becomes:

$$\begin{aligned}&-\nu_1(a-x)+\tfrac{3}{2}\nu_2[-2a^2x+3ax^2+ay^2-x(x^2+y^2)]\\&-\frac{15}{8}\nu_3[4a^3x^2-8a^2x^3-4a^2xy^2+4ax^2(x^2+y^2)\\&+a(x^2+y^2)^2]+\frac{35}{16}\nu_4[-8a^4x^3+8a^3x^4\\&+12a^3x^2(x^2+y^2)]-\frac{315}{128}\nu_5.16a^5x^4 \quad (24)\end{aligned}$$

where

$$\nu_n = q\delta a/(a^2+z^2)^{n+\frac{1}{2}},\ q = \Phi_s\rho$$

The effect of these "displacement terms" on the motion of an electron will now be considered.

The general ray-equation may be written:

$$(\partial/\partial x - d/dz\partial/\partial x')\sqrt{[\phi(1+x'^2+y'^2)]} = 0 \quad (25)$$

$$(\partial/\partial y - d/dz\partial/\partial y')\sqrt{[\phi(1+x'^2+y'^2)]} = 0 \quad (26)$$

where primes denote differentiation with respect to z. (This is the non-relativistic form; inclusion of relativistic terms would add to complication without greatly affecting the conclusions.) This yields[5]:

$$(\partial\phi/\partial x - x'\partial\phi/\partial z)(1+x'^2+y'^2) = 2\phi x'' \quad (27)$$

$$(\partial\phi/\partial y - y'\partial\phi/\partial z)(1+x'^2+y'^2) = 2\phi y'' \quad (28)$$

For a system containing merely a round lens and Φ_2 and Φ_4 units, the expression for the potential is:

$$\phi = \Phi - \frac{\Phi''}{4}(x^2+y^2)+\Phi_2(x^2-y^2)+\frac{\Phi''''}{64}(x^2+y^2)^2+\frac{\Phi_2''}{12}(y^4-x^4)+\Phi_4(x^4-6x^2y^2+y^4)+0(x^6) \quad (29)$$

[see, for example, equations (1.2) and (3.1) in Ref. 1]. But now add to this the difference terms expression (24) (which were due to the geometrical displacement) and apply equation (27); there results:

$$\begin{aligned}&x''+x'\frac{\Phi'}{2\Phi}+x\frac{\Phi''}{4\Phi}\\&=\Big[x'y'y''-x''y'^2+\frac{1}{4}\frac{\Phi''}{\Phi}x''(x^2+y^2)\\&+\frac{1}{32}\frac{\Phi''''}{\Phi}x(x^2+y^2)-\frac{1}{2}\frac{\Phi''}{\Phi}xy'^2-\frac{1}{4}\frac{\Phi''}{\Phi}xx'^2\\&+\frac{1}{8}\frac{\Phi'''}{\Phi}x'(x^2+y^2)-\frac{1}{2}\frac{\Phi'}{\Phi}x'(x'^2+y'^2)\\&+\frac{1}{4}\frac{\Phi''}{\Phi}x'yy'\Big]+\Big[\frac{\Phi_2}{\Phi}x-\frac{\Phi_2}{\Phi}x''(x^2-y^2)\\&-\frac{1}{6}\frac{\Phi_2''}{\Phi}x^3+2\frac{\Phi_2}{\Phi}xy'^2-\frac{\Phi_2'}{2\Phi}x'(x^2-y^2)+\frac{\Phi_2}{\Phi}x'^2x\\&+\frac{\Phi_2}{\Phi}x'yy'+\frac{\Phi_4}{\Phi}2x(x^2-3y^2)\Big]\\&+\Big\{\Big[\nu_1+\frac{3}{2}\nu_2(-2a^2+6ax-3x^2-y^2)\\&-\frac{15}{8}\nu_3(8a^3x-24a^2x^2-4a^2y^2+20ax^3+12axy^2)\\&+\frac{35}{16}\nu_4(-24a^4x^2+80a^3x^3+24a^3xy^2)\\&-\frac{315}{2}\nu_5a^5x^3\Big](1+x'^2+y'^2)-x'\Big[-\nu_1'(a-x)\\&+\frac{3}{2}\nu_2'(-2a^2x+3ax^2+ay^2)-\frac{15}{2}\nu_3'a^3x^2\Big]\\&+a\nu_1'x'(x'^2+y'^2)-2x''[-\nu_1(a-x)\\&+\frac{3}{2}\nu_2(-2a^2x+3ax^2+ay^2)-\frac{15}{2}\nu_3a^3x^2]\Big\}/2\Phi \quad (30)\end{aligned}$$

If equation (30) be written $L(x) = F(x, y)$ and x_α, x_γ are two independent solutions of $L(x) = 0$, then it may be shown (e.g. Burfoot[6]) that the solution of equation (30) takes the form:

$$x = \alpha x_\alpha + x_0 x_\gamma + \frac{x_\alpha}{\sqrt{\Phi_A}}\int_{z_0}^{z}\sqrt{\Phi}.Fx_\gamma dz - \frac{x_\gamma}{\sqrt{\Phi_A}}\int_{z_0}^{z}\sqrt{\Phi}.Fx_\alpha dz \quad (31)$$

where $x_{\alpha 0} = 0$, $x'_{\alpha 0} = 1$, $x_{\gamma 0} = 1$.

As it stands, equation (30) is in the form used by Burfoot; this has certain computational advantages over that of Scherzer, in which the $(\Phi_2/\Phi)x$ was included in the $L(x)$. It will, nevertheless, be convenient for present purposes to use Scherzer's form, and to go a little beyond it by transferring to the left-hand side the terms:

$$(9\nu_2 ax - 15\nu_3 a^3 x)/2\Phi \quad (32)$$

These contain x to the *first* order, as does $(\Phi_2/\Phi)x$, but they are *small* because the ν's contain δa.

The arrangement of the correction systems described in Part 1 have all been such that, in the absence of the displacement terms, x_α and y_α vanish together in the image plane. But plainly the displacement terms will not have the same

effect on the "y" equation corresponding to equation (30) [i.e. the equation resulting from equations (24) and (28)], since the assumed displacement was in the x-sense. Consequently the zeros of x_α and y_α will move apart, i.e. there will be introduced a first order astigmatism. This does not matter, however, as a means of dealing with it is already available, namely the Stigmator (a rotatable Φ_2 unit—Rang[4]) which in any case will normally be needed in addition to a spherical correction system in order to allow for mechanical inaccuracies in the construction of the lens under correction.

Thus, aberrations in the image plane will be obtained (i) from the last term of equation (31), in which F now comprises the right-hand side of equation (30), less the terms mentioned, and (ii) from the corresponding "y" equation; (i) will now be considered alone and (ii) restored later. It has already been pointed out that x_{yi} (i : image position) = y_{yi} = magnification M. The third order terms contained in the first square bracket of equation (30) constitute the ordinary spherical aberration of a round lens (by partial integrations they may be proved to be equal to Scherzer's spherical aberration expression). Those in the second square bracket comprise the correction terms due to the introduction of the Φ_2 and Φ_4 units. It is the principle of correction systems to make these cancel. The zero order terms $(\nu_1 - 3a^2\nu_2)/2\Phi$ represent a small constant displacement which does not affect aberrations. There remain terms of the second and third degree in x, y.

Consider rays emerging from the axial point of the object ($x_0 = 0$). Then in the small terms comprised in F it is appropriate to put: $x = \alpha x_\alpha$, $y = \beta y_\alpha$; where α, β are the starting angles in the xz, yz planes. As these may normally be expected to be of order 1/100, it is reasonable at first to neglect the third order terms of equation (30) and consider only the second. These are:

$$\Big[(\nu_1 - 3a^2\nu_2)(\alpha^2 x_\alpha'^2 + \beta^2 y_\alpha'^2) - \alpha^2 x_\alpha x_\alpha' \nu_1' + 3\alpha^2 x_\alpha x_\alpha' \nu_2' a^2 - 2\alpha^2 x_\alpha x_\alpha'' (\nu_1 - 3a^2\nu_2) - \frac{3}{2}\nu_2(3\alpha^2 x_\alpha^2 + \beta^2 y_\alpha^2) + \frac{15}{8}\nu_3(24a^2\alpha^2 x_\alpha^2 + 4a^2\beta^2 y_\alpha^2) - \frac{105}{2}\nu_4 a^4 \alpha^2 x_\alpha^2\Big]/2\Phi \quad (33)$$

Now in the region where these displacement terms arise, x_α and/or y_α are of the order of f, the focal length of the lens under correction. Also a may be expected to be of order f, whereas x_α', y_α' are of order $1/M$. It appears therefore that the terms without dashes in equation (33) predominate. Taking x_α, y_α constant over the unit, as before, integration yields:

$$M\frac{6}{a^4}(\beta^2 - \alpha^2) f^3 \Phi_s \delta a . \rho / 2\Phi_A \quad (34)$$

α, β are independent, so this is greatest when one of them is zero. Suppose this residual aberration is not to exceed $MC_s\alpha^3/n$. Then tolerance δa is given by:

$$\delta a < \frac{C_s . \alpha}{n} \frac{2\Phi_A}{\Phi_s} \frac{a^4}{6f^3\rho} \quad (35)$$

If the expression for the potential on one of the pair of Φ_4 units (17) is inserted, the tolerance becomes:

$$\delta a < 2f\alpha/n \quad (36)$$

Resolution (as limited by the combined action of diffraction and spherical aberration) improves inversely as the fourth root of the spherical aberration; hence a resolution improvement of three times (which would not be without value) would require $n = 81$; with $f = 0{\cdot}5$ cm, $\alpha = 0{\cdot}01$, this would require a tolerance δa of $1{\cdot}2\,\mu$. It would be desirable to increase both n and δa, and it remains to be seen whether any device can be arranged to correct the second order aberrations in the same way as Rang's Stigmator corrected the first.

Removal of second order aberrations

Consider the effect of a unit comprising *two* spheres [Fig. 1(*c*)]. An expression for its potential at (x, y, z) has already been given in expression (12). If equation (27) be applied to the combination of this with previous units, the effect is to add to the right-hand side of equation (30) the terms:

$$[2a\nu_1 - 3a\nu_2(3x^2 + y^2) + 15a^3\nu_3 x^2 + 2a\nu_1(x'^2 + y'^2) - 2a\nu_1' xx' - 4a\nu_1 xx'']/(2\Phi_A \delta a) \quad (37)$$

When x, y are put equal to αx_α, βy_α respectively, this provides terms in α^2, β^2, not necessarily in the same ratio as the α^2, β^2 terms of expression (33). Again, if the companion equation of equation (30) [derived from equations (24) and (28)] be written out, it will be found to contain terms in $\alpha\beta$; if equation (28) now be applied to the extra unit (37), further terms in $\alpha\beta$ will result, but these will not, in general, bear such a relation to the α^2, β^2 terms as to permit complete second order correction by means of *one* two-sphere unit. It will, in fact, prove simplest to use three such "Φ_1" units, located at the positions occupied by the original Φ_4 units of Seeliger, namely, one at the zero of x_α (which provides β^2 terms but not α^2, $\alpha\beta$), one at the zero of y_α (which provides α^2 terms but not $\alpha\beta$, β^2), and one where x_α and y_α are both finite (providing α^2, $\alpha\beta$, β^2).

So far, consideration has been restricted to the effect of a radial displacement of one of the "x" spheres (Fig. 1). But analogous calculations show that a radial displacement of one of the "y" spheres would require a Φ_1 unit orientated at $\pi/2$ to the former; the two could be combined to form a single Φ_1 unit orientated in an intermediate direction. Thus, *three* separate Φ_1 units would suffice to account for the α^2, $\alpha\beta$, β^2 terms produced by radial displacements of *any* of the spheres, provided that the units were made capable of rotation (compare Rang's Stigmator).

The addition of three extra rotatable units would considerably complicate a correcting system, but the same effect could be obtained by arranging that the potentials on any of the four pairs of opposite spheres of the Φ_4 units might be varied a little *anti-symmetrically*. If none of the four thus available directions coincided with that appropriate to the effective Φ_1 unit, then the two pairs enclosing the desired direction might be adjusted so that their resultant did so. A means of observing the Φ_1 effects will be considered later.

Fourth order aberrations

Let it now be assumed that a Φ_1 system has been included and that second order aberrations have thereby been removed. The residual displacement terms of equation (30) are all of the third order. But these may be absorbed into the already existing third order effect of the Φ_4's. In other words, they may be regarded as a little extra spherical aberration which may be removed by slightly varying the potential on the Φ_4 units. In order now to determine tolerances, it is necessary to pass to the *fourth* order. This adds to equation (24):

$$(15/8)\nu_3 r^5 \cos\theta - (35/16)\nu_4(6a^2r^5\cos\theta + 12a^2r^5\cos^3\theta) + (315/128)\nu_5(32a^4r^5\cos^3\theta + 16a^4r^5\cos^5\theta) - (693/8)\nu_6 a^6 r^5\cos^5\theta \quad (38)$$

and after due application of equations (27) and (31) yields an aberration of:

$$M\frac{f^5}{a^6}[10(\alpha^4+\beta^4) - 60\alpha^2\beta^2]\frac{\Phi_s\cdot\rho\delta a}{2\Phi_A} \quad (39)$$

This will be greatest when $\alpha = \beta$, and tolerance will therefore be:

$$\delta a < \frac{C_s}{\alpha n}\frac{\Phi_A}{\Phi_s}\frac{a^6}{f^5}\frac{1}{20\rho} \quad (40)$$

or, employing equation (17) (for $\Phi_{s(4p)}$):

$$\delta a < 0{\cdot}3a^2/\alpha n f \quad (41)$$

Using again $a = 0{\cdot}4$ cm, $f = 0{\cdot}5$ cm, $\alpha = 0{\cdot}01$, but now increasing n to 10^4 so as to correspond to a resolution improvement factor of 10, this gives $\delta a = 10\ \mu$, or about 5×10^{-4} in.

This represents the tolerance for the pair of Φ_4 units in Seeliger's system; that of the third Φ_4 unit is four times less stringent. For the Φ_2 units, equation (22) may be used: with $M = 20$, $C_s = 0{\cdot}5$ cm, $\rho = 0{\cdot}1$ cm as before, tolerance ($n = 10^4$) is $\delta a = 40\ \mu$.

Effect of displacing one sphere tangentially

If one sphere is rotated $\delta\theta$ about the axis, there is a change of potential at (r, θ, z):

$$\frac{q}{(a^2+z^2)^{\frac{1}{2}}}\left\{\left[1 + \frac{r^2 - 2ar\cos\theta + 2ar\sin\theta\delta\theta}{(a^2+z^2)}\right]^{-\frac{1}{2}} - \left[1 + \frac{r^2 - 2ar\cos\theta}{(a^2+z^2)}\right]^{-\frac{1}{2}}\right\} \quad (42)$$

Changing to x, y and limiting to fifth order terms, this gives:

$$\left\{-\nu_1 y + \frac{3}{2}\nu_2 y(x^2+y^2-2ax) - \frac{15}{8}\nu_3 y[(x^2+y^2)^2 - 4ax(x^2+y^2) + 4a^2x^2] + \frac{35}{16}\nu_4 y[12a^2x^2(x^2+y^2) - 8a^3x^3] - \frac{315}{8}\nu_5 y a^4 x^4\right\}(a\delta\theta/\delta a) \quad (43)$$

These terms are of the same nature as expression (24) and its "y" companion: on the assumption that the Stigmator and Φ_1 system have been duly incorporated, tolerances will come from the application of equation (28) [here more important than equation (27), as the shift is in a "y" sense] to the fifth order in expression (43). The leading terms of the resulting aberration are:

$$M\frac{f^5}{a^6}[10(\alpha^4+\beta^4) - 60\alpha^2\beta^2]\frac{\Phi_s\rho a\delta\theta}{2\Phi_A} \quad (44)$$

[compare expression (39)]. Thus, the tangential tolerance $a\delta\theta$ is precisely the same as the radial tolerance δa given in expression (39).

Detection of second order aberrations

First order astigmatism is revealed as the occurrence of two separate "image" positions, at which objects lying perpendicularly to one another are respectively in sharp focus. Second order aberrations merely blur these part-images to different degrees. It would be hard to estimate when the part-images were equally well focused and hence second order aberrations absent.

Regenstreif[7] has given the theory (experimentally justified) of the caustics resulting from various forms of asymmetry. In particular, that due to first order astigmatism takes the form of a four-pointed star. The addition of asymmetries corresponding to second order aberrations (Φ_1 and Φ_3) adds more points to the star, but if the extra aberration is small the general effect is an anti-symmetrical change in the lengths of opposite arms. This may prove to be a better means of detecting odd asymmetries; if the Φ_1 units incorporated in a correcting system are so adjusted as to make the caustic four-symmetrical, it is clear that only Φ_2, Φ_4 effects remain, and these can subsequently be adjusted to achieve spherical and/or chromatic correction.

CONCLUSIONS

Part 1. A modification of Seeliger's spherical correction system would remove the need for cylindrical lenses at the expense of adding one round lens; if this extra lens were regarded as a part of the lens under correction, the system would be analogous to that of Burfoot. The place of the cylindrical lenses would be taken by Φ_2 units (having the same form as Rang's Stigmator).

The potentials required for spherical-aberration-correcting units (Φ_2 and Φ_4) comprising arrays of spheres round the axis would range from ± 2 to 16 kV (for overall diameter of unit 1 cm) [equations (15) and (16)], but this could conveniently be reduced by a factor of 20 by using laminar electrodes (as in Seeliger's apparatus) instead of spheres.

Part 2. Expressions have been given for the positioning tolerances of Φ_2 and Φ_4 units comprising arrays of spheres [expressions (35) and (40)]. For a three-times improvement in resolution the spheres would have to be positioned accurately to 1 μ. With the addition of three rotatable Φ_1 units (pairs of oppositely charged spheres) to the correction system, the improvement in resolution could be increased to a factor of ten times and the tolerances increased to 10 μ. The complication introduced by the need for extra units could be avoided by providing a slight independent variation of the potentials on individual spheres of the Φ_4 units. This would resemble the practical arrangement for adjustment of the stigmator already used by Haine and Mulvey in this Laboratory, concerning which a paper will appear later.[8]

ACKNOWLEDGEMENTS

The author wishes to thank Dr. V. E. Cosslett and Dr. J. C. Burfoot for permitting the study of material before publication, Mr. M. E. Haine and Dr. G. Liebmann for most helpful comment, and Dr. T. E. Allibone for allowing the publication of this paper.

REFERENCES

(1) Scherzer, O. *Optik*, **2**, pp. 114–32 (1947).
(2) Seeliger, R. *Optik*, **8**, pp. 311–17 (1951).
(3) Sturrock, P. A. *Phil. Trans. A*, **243**, pp. 387–429 (1951).
(4) Rang, O. *Optik*, **5**, pp. 518–30 (1949).
(5) Burfoot, J. C. *Proc. Phys. Soc.* [*London*] *B*, **66**, pp. 775–92, 1953.
(6) Burfoot, J. C. Thesis, University of Cambridge (1952).
(7) Regenstreif, E. *Ann. Radioélect*, **6**, pp. 51–83 (1951).
(8) Haine, M. E., and Mulvey, T. *J. Sci. Instrum.*, to be published (1954).

New books

High altitude rocket research. By HOMER E. NEWELL, JR. (New York: Academic Press Inc.) Pp. 298. Price $7.50.

This volume is an up-to-date progress report on American rocket sounding of the higher atmosphere, carried out at the Army Ordnance White Sands Proving Ground and on the U.S.S. *Norton Sound*, a naval vessel assigned to guided-missile research. Its author is the head of the Rocket-Sonde Research Branch of the Naval Research Laboratory, Washington, and so is uniquely qualified for his task. Experiments using V2 rockets and their descendants, the Viking and the Aerobee, have, since 1946, extended the direct-measurement approach to atmospheric characteristics, beyond the balloon ceiling of about 30 kilometres to almost 400 kilometres. Information secured in flight is either sent back by radio or instrumentally stored to be available in the rocket débris on arrival at the ground.

It is a fortunate fact that the ambient atmospheric pressure can be measured directly, during flight, at a place just ahead of the rocket tail fins; and, once the pressure-height curve is available, corresponding curves for atmospheric temperature and density can be deduced. Such work has confirmed the existence of the warm stratum of air at 50 kilometres, postulated earlier by British workers from indirect evidence. Still more recent rocket soundings have reached the ionosphere, where measurements of both electron density and geomagnetic field have been made.

The book is not overburdened with technical detail, but the scientific basis of every type of investigation is lucidly and concisely explained. It is just the nature of the situation that some of the results described will have been substantially extended, or even corrected, by work conducted while the book was printing.

EDWARD V. APPLETON

Television receiver design, Monograph 2. Flywheel synchronization of saw-tooth generators. By P. A. NEETESON. (London: Cleaver Hume Press Co. Ltd.) Pp. 156. Price 21*s*.

This book, as the title indicates, is one of a series on television receiver design. It is an excellently produced book and a useful addition to the library of a well-equipped laboratory. A criticism of this book can perhaps be summarized by criticizing its title since a good deal of material bordering on the subject is included whilst the scope on flywheel time-base circuits themselves, particularly on the practical side, is disappointingly narrow. As examples of this the first one-third of the book is a review of the general principles of time-base circuits with a description of the special type valves involved, and some 22 pages are devoted to the flywheel action of tuned circuits. The description of practical circuits is limited to the last 10 pages. The general attack on the analytical side is very good. The way in which the frequency of multivibrator and blocking oscillator circuits can be controlled by change of operating conditions is theoretically examined and experimental results on actual circuits are given to substantiate the theory. Various arrangements are then studied by graphical methods to show the operating characteristics of complete automatic phase control systems. The influence of various disturbances and the effects of filtering circuits in the control path are included. The book will appeal to those who have to tackle the important problems of flywheel synchronization in a really fundamental way.

A. J. BIGGS

The electrical breakdown of gases. By PROFESSOR J. M. MEEK, D.Eng., F.Inst.P., and J. D. CRAGGS, M.Sc., Ph.D., F.Inst.P. (London: Oxford University Press.) Pp. vii + 507. Price 60*s*.

The widespread ramifications of the subject of the electrical breakdown of gases, both in the field of electronics and in power electrical engineering, make it unnecessary to stress the interest which a book such as this will arouse. The appearance of a comprehensive treatise on the subject is therefore to be welcomed, especially since the authors, Professor Meek and Dr. Craggs, have contributed much to the literature of the subject. Chapter headings indicate the scope of this treatise: fundamental processes in electrical discharges; breakdown at low gas pressures; corona discharges; experimental studies of the growth of spark discharges; the lightning discharge; theory of the spark mechanism; breakdown voltage characteristics; irradiation and time lags; high frequency breakdown; the spark channel; electrode effects; and glow to arc transitions. It is impossible in a brief review to survey each chapter in detail.

In their preface the authors state that they have confined themselves largely to a review of papers published since 1930, but the appearance of numerous papers in very recent years has necessitated recourse to a substantial appendix. The text itself is somewhat uncritical from the point of view of assessing the present position, and the reader may encounter some difficulty when finding in the appendix statements and conclusions which not only modify, but sometimes are at variance with the text. The book gives, with its extensive and valuable list of references, an extremely comprehensive account of work done during the last twenty years.

However, some chapters give a large number of observations which, from their very complexity, make it difficult to reach conclusions of a fundamental nature. For instance, Chapter 3 on corona discharges hardly gives an adequate discussion of breakdown under non-uniform fields, and clarity is not served by the use of confusing expressions such as: *burst pulse onset* or *pre-onset burst pulse corona*. Again, Chapter 4, on the growth of the spark discharge, contains a wealth of diagrams and cloud chamber photographs, but it is not easy to see how this helps a basic understanding of the subject. This is a field in which simple qualitative views can be misleading, and the interpretation of much of this data is impossible. Here the reader feels the need of a fundamental mathematical treatment of the spatial-temporal growth of ionization.

The theory of the spark mechanism is given in Chapters 2 and 6, possibly the two most important chapters in the book. Chapter 2 discusses breakdown at the lower gas pressures, and gives the Townsend theory of current growth dependent on primary and secondary ionization processes. Chapter 6 then opens with the statement that this theory encounters difficulties for example in connexion with the speed of breakdown under impulse conditions. Unfortunately, no critical analysis is made to show whether experimental data lead to this conclusion, and the reader cannot see exactly how the

supposed difficulties arise. Such an analysis is crucial to any understanding of the physical significance of breakdown under impulse conditions. (There would be less confusion here if there were more co-ordination of text and appendix.) Chapter 6 then goes on to describe a different theory of the spark mechanism (the streamer theory). Propounded in 1940 in the *Mechanism of the electric spark* by Professor Loeb and the senior author of the present book, this theory was based on a number of unconfirmed assumptions relating to the action of space charges and the effects of photo-ionization. It is somewhat surprising to find in the present book a statement of this theory in the same qualitative empirical terms as those originally used: apparently no theoretical or experimental confirmation has been established in the intervening 13 years, and the range of application claimed for the theory is now still further restricted.

Although they have unhappily not taken the opportunity of attempting to assess how much of the subject is now amenable to a systematic treatment, the authors are nevertheless to be congratulated on producing this most comprehensive survey of an enormous field. This book is likely to be regarded in research laboratories as a very valuable and helpful work of reference, the need for which has long been felt. The production of the book, diagrams and photographs have the excellence associated with The Clarendon Press.

F. Llewellyn Jones

Metallurgy of the rare metals. No. 1. **Chromium.** By A. H. Sully, M.Sc., Ph.D., F.I.M., F.Inst.P. Pp. xii + 272. Price 35*s*., postage 1*s*. extra. No. 2. **Zirconium.** By G. L. Miller, B.Sc., Ph.D., A.R.I.C., M.I.Chem.E. Pp. xviii + 382. Price 45*s*., postage 1*s*. 6*d*. extra. (London: Butterworths Scientific Publications Ltd.; New York: Academic Press Inc.)

These two volumes, on chromium and zirconium, form the first of a series of surveys which is to include titanium, molybdenum, the platinum metals, manganese and uranium, appearing under the omnibus title *The metallurgy of the rarer metals*. While the title of the series is rather misleading in its effort to be brief, it is not difficult to see what has been in the mind of its general editor, Dr. H. M. Finniston, who is head of the Metallurgy Division at the Atomic Energy Research Establishment, Harwell. There has been, in recent years, an intense effort to prepare these metals more cheaply and in purer forms than hitherto for the wide technological promise offered by them and by alloys based on the pure metals, and it is opportune that the considerable accumulation of new knowledge should now be summarized and weighed up.

Since much of the stimulus to study and use these "rarer" metals stems from the activities of physicists it is not inappropriate that the first volume of the series should be written by a physicist. Dr. Sully, in his book on chromium, subdues his inclinations towards "metal physics," however, and presents a broad review of his subject which must rank as the definitive text on chromium.

Chromium in combined form is not particularly rare in nature and indeed there are many rich sources of chromium ores. The metal itself has been widely used for many years alloyed with nickel and as a constituent of alloy steels and, more recently in the form of surface coatings applied by diffusion and by electrodeposition, the familiar chromium plate.

The valuable hardening and oxidation-resisting properties conferred by the metal when used in these ways raised hopes of the development of other, chromium-rich alloys having similar desirable properties. However, electrolytic chromium, which can be readily prepared substantially free from metallic impurities and can subsequently be treated to reduce the substantial oxygen content to a low figure, has hitherto always been found to be brittle at normal temperatures. This is true also at elevated temperatures of the chromium-base alloys which have been developed having outstanding creep and oxidation resistance at temperatures of the order of 900° C. Dr. Sully concludes on this account that the potentialities of chromium-base alloys as engineering materials are not at all good, and the possibility of ameliorating their disadvantages remote. Nevertheless, he refers to recent work from which it appears that chromium may in some circumstances be quite ductile at room temperature, which perhaps leaves room for a little hope that his gloomy prediction may be wrong.

Like Dr. Sully, Dr. Miller effectively surveys the broad field of information on his subject, dealing with occurrence, extraction, physical properties, alloy systems and industrial uses. Again the book fills a gap in timely fashion. Zirconium is particularly attractive on account of its resistance to corrosion by liquids, which, coupled with a relatively high melting point and low absorption for slow neutrons, suggests applications in the field of atomic energy engineering, as well as in the broader chemical engineering field. It has special applications in surgery and as a getter in vacuum electronic devices, and has been used for a number of years as a minor, but important, constituent of magnesium and copper alloys, and as a scavenger in steel production.

Like chromium, zirconium is a highly reactive metal, taking up the common gases and readily attacking refractories. The chapters on reactions with gases and on commercial production methods for melting and fabricating the metal and its alloys are thus of particular interest and value. As a useful by-product of its reactivity, zirconium provides a practical means of joining other metals to refractory non-metallics like diamond, carbides and oxides, and it forms hard, high melting point, metallic compounds such as borides and carbides which may have useful high-temperature engineering applications. These topics are also well surveyed.

Work on both chromium and zirconium continues, and it will probably not be long before there is important new information to be reviewed. It is gratifying to note, therefore, that the publishers aim to include the new information available as successive editions go to press.

G. L. J. Bailey

Progress in cosmic ray physics, Vol. 2. Edited by J. G. Wilson. (Amsterdam: North-Holland Publishing Co.) Pp. xi + 322. Price £3.

Professor J. G. Wilson and the publishers are to be congratulated on producing the second volume of review articles on progress in cosmic ray physics. The first volume has proved to be very useful and it is to be hoped that a volume will appear each year. Volume 2 contains five articles written by acknowledged experts and is excellently produced.

R. D. Sard and M. F. Crouch give an authoritative account of the nuclear interactions of stopped μ-mesons, a subject which is in an advanced state of development, and this article will undoubtedly remain the standard work of reference for some time. The second article, by J. G. Wilson, on experimental data on the heavy unstable particles is an excellent and comprehensive review of the position in this field towards the end of 1952. The tempo of research work on the properties of the heavy unstable particles has been gradually increasing since their discovery in 1947, but nevertheless the amount of

reliable data is still meagre. It is not surprising, therefore, that some of the conclusions put forward tentatively by Wilson cannot be regarded as correct at the beginning of 1954. There is no doubt, however, that review articles on the new particles are of considerable value to young researchers, particularly nuclear physicists who may wish to work with artificial heavy particles. Unfortunately, they all date very rapidly and thus it is very doubtful if the Editor would be wise to include such an article as a regular feature of his series unless arrangements can be made for speedy publication or considerable revision in proof. The late E. G. Dymond has written a short account of the experiments on the intensity of the penetrating component of cosmic radiation in the upper atmosphere consisting of primary particles and secondary mesons. The energetic primary radiation interacts readily with the earth's atmosphere, producing complex nucleon and electron-photon cascades. Extensive theoretical studies of these cascades have been made by H. Messel and his colleagues during recent years. Messel has now written a long review of this work which will be a considerable help to experimentalists, many of whom have great difficulty in understanding the many complex theoretical papers. The last article by L. Voyvodic deals with the techniques of particle identification using photographic emulsions. A summary is given of the main properties of nuclear emulsions, followed by a detailed discussion of the available methods for the measurement of particle mass and energy. C. C. BUTLER

An introduction to the theory of seismology. By K. E. BULLEN. (London: Cambridge University Press.) Pp. xv + 296. Price 35*s*.

This is the second edition of a most valued and comprehensive text which was first published in 1947. This edition not only brings up to date the information on recent developments, but is also much improved by the inclusion of a twenty-page list of references grouped according to subject matter. The author treats his subject, generally speaking, from the mathematical point of view. While this is an essential feature for the serious student, it inevitably makes the book, save for three or four chapters, unsuitable for those who are not equipped beyond the stage of elementary mathematics. The whole field of knowledge of seismology is covered in an authoritative way, although, of course, Professor Bullen's close association with the development of earthquake seismology on the global scale, results in this branch receiving the most attention. However, those who are mainly interested in the small-scale phenomena of seismic prospecting will find much that is both interesting and applicable to that sphere.

L. H. TARRANT

The determination of crystal structures. By H. LIPSON and W. COCHRAN. (London: G. Bell and Co. Ltd.) Pp. ix + 345. Price 50*s*.

This book, used in conjunction with the *International tables for X-ray crystallography*, is undoubtedly Everyman's guide to crystal structure determination: it is a worthy companion to the first two volumes of *The crystalline state* and should be warmly welcomed. It deals comprehensively with X-ray structure analysis from the stage at which the structure amplitudes have been obtained from the experimental measurements to the final determination of the atomic positions. The authors have done excellent work in sifting and abstracting the original literature to provide complete accounts of every important aspect of this long and intricate process, and they have made the subject live by means of a wealth of examples, among which they have included both elementary and complicated cases, in pursuance of their aim to make the book useful both to the beginner and to the more advanced worker.

As the authors point out, the fundamental principles employed in structure analysis are essentially those of physical optics, and Volume 2 of this series, by R. W. James, was devoted to the optical principles of X-ray diffraction. However, while the principles remain the same, emphasis on different ways of applying them is apt to shift and this is illustrated by the importance rightly assigned in the present volume to Fourier transforms, which merited no more than a footnote from Professor James in 1948. The account of this topic and of the optical methods especially associated with Dr. Lipson's name is a particularly welcome feature of the book.

Every critical reader of a book dealing with a rapidly developing subject such as this is bound to encounter expressions of opinion with which he disagrees, and the reviewer hopes that the view (p. 106) that electron counts are easy to make in three dimensions and ought to be made more often will not encourage the making of too many unreliable estimates. There are serious difficulties in the way of making accurate electron counts and probably for the majority of structure determinations at the present time the approximate check provided by peak heights is all that is justified.

The book is well produced and copiously illustrated (with the same convenient system of numbering the diagrams which was used in Henry, Lipson and Wooster's *Interpretation of X-ray diffraction photographs*); having regard to the general rise in costs since 1948 its price is reasonable compared with that of Volume 2 of this series.

E. G. Cox

Applied elasticity. By CHI-TEH WANG, Sc.D. (London: McGraw-Hill Publishing Co. Ltd.) Pp. ix + 357. Price 57*s*. 6*d*.

Any new book on elasticity will be judged mainly on its presentation, and on this basis this book will be welcomed especially by those who still find theoretical elasticity a baffling mystery. The subject is treated in a lucid and readable style which often surpasses in quality the matter with which it deals.

The treatment of plane strain, plane stress and generalized plain stress reflects the confusion that now exists especially in American literature. Anyone trying to use this chapter as a text book will be left with unresolved discrepancies. The Lamé definition of shear strain, containing the coefficient $\frac{1}{2}$ is more logical than that used by the author. Valuable analogies between stress and strain are lost by not using it. The derivation of the stress-strain relations would have greater appeal to "applied elasticians" had it been obtained using physical concepts of strain energy, and isotropy. The relation for non-isotropic material could then have been obtained with complete relevance as well as the proof that principal directions of stress and strain coincide. In the proof of the uniqueness theorem the sentence "if the principle of superposition can be applied" is, at best, meaningless. It is surely bad policy to throw doubt on so extensively used a principle without discussing more fully its limitations. In these early chapters the assumptions required for a given result should have been more clearly stated. For example, the equation of equilibrium, not limited to elastic theory, is applicable under plastic conditions. This is easily understood when the necessary assumptions are clearly set out.

Of the author's two aims, to teach fundamentals, and to leave the student able to solve practical problems, the second is more completely realized. This part of the book is convincing and covers a wide field. Due mainly to space limitation the chapter on complex variable fails to convince one of the greater power and usefulness of this method.

The restriction to "elementary mathematics" is an advantage and the occasional loss of rigour is due mainly to incomplete argument. C. SNELL

Relativity—the special and the general theory. By ALBERT EINSTEIN, translated by R. W. LAWSON. (London: Methuen and Co. Ltd.) Pp. x + 165. Price 12*s.* 6*d.*

This book of Einstein's was first published in 1920, at a time when relativity excited scientific and non-scientific circles in much the same way as atomic energy does today. Of course, the two subjects are closely related, for one is a direct descendant of the other. To the present reviewer, who found an early edition of the book very difficult going nearly thirty years ago, the new fifteenth edition appears as clear and lucid as can be achieved in such small compass. As the book has changed little over the years, he must conclude that amongst other things he had not the "patience and force of will" which Einstein, in the preface, confesses the reader needs, adding that "the work presumes a standard of education corresponding to that of a University matriculation examination"! This is a necessary but hardly a sufficient condition.

The book consists of many short sections in which the main ideas of both the special and the general theories are expounded without the use of advanced mathematics, or of any mathematics at all. There are a few appendices which include a simple derivation of the Lorentz transformation and some account of the experimental confirmation of the general theory. The feature of this edition is the addition of Appendix V in which Einstein presents his most recent views on "space-time".

The translator has added a good bibliography.

J. TOPPING

Principles of transistor circuits. Edited by R. F. SHEA. (London: Chapman and Hall Ltd.; New York: John Wiley and Sons, Inc.) Pp. xxx + 535. Price 88*s.*

The rapid development of the transistor since its inception six years ago has made a comprehensive treatment of the principles of its operation a matter of some urgency to ensure the full utilization of its potentialities. The authors of this book, who have taken an active part in the development of transistor circuits at the General Electric Laboratories, have now produced a very comprehensive compendium on all aspects of transistor circuitry which should prove extremely useful to anyone having to design such circuits. Even the builder of the transistor itself might find some useful hints on ways of designing a transistor to a given specification. The student approaching the subject for the first time might, however, find the going rather difficult at times.

The book is remarkably up to date, but it is an inherent danger with any book on a subject still in the full flux of development that it might soon become out of date. The authors have obviously been aware of this and tried to avoid the danger by stressing general principles. They first discuss, briefly, the physics of semiconductors and of transistor action, then describe the various types of transistors and establish equivalent circuits before considering in detail the application to various types of amplifiers and oscillators. There are useful chapters on high-frequency operation, noise, transient response and computor circuits, as well as on the application of matrix methods and on design by the principle of duality.

It is to be hoped that the book will be kept up to date by revising it as the transistor reaches maturity. So rapid is the present development that certain aspects would already need amplification, as for instances the use of alternative materials to germanium, especially silicon and certain binary semiconducting compounds, a discussion of the surface barrier transistor and of the potentialities of combining *p–n–p* with *n–p–n* transistors, etc.

The production of the book is excellent, the type and format are well chosen, there are many useful illustrations and examples, a detailed subject index and a comprehensive bibliography up to the end of 1952. E. BILLIG

Structure reports for 1945–46, Vol. 10. Edited by A. J. C. WILSON. (Utrecht: A. Oosthoek Publishing Co.) Pp. viii + 325. Price 45*s.*

This volume is one of the three "back numbers" planned to bridge the gap between the last issue of *Strukturbericht* (Volume 7 for 1939) and Volume 11 of *Structure reports* (for 1947–48). It deals with papers of structural interest published in the years 1945–46, excluding a few already mentioned in Volumes 11 and 12. The three sections, on metals, inorganic structures, and organic structures, respectively, each occupy about 100 pages.

The Editors state that it is intended to report only material of structural interest and to do this so completely that no further structural information would be gained by consulting the original paper. It is, however, somewhat difficult to define the limits of structural interest, with the result that some reports, as for instance on "Fractographic structures in bismuth," might not be generally considered suitable for inclusion. The latter aim is extremely difficult to fulfil in reasonable space for long and complicated papers, and some reports on such papers are perhaps lacking in clarity, as for instance that on chrysotile and serpentine (page 160).

The arrangement within each report follows the excellent scheme laid down in previous volumes, and, as before, separate indexes for subjects, authors, and for formulae make it easy to locate any required report. Editorial comments, clearly indicated by square brackets, are in general very helpful even if occasionally trivial as on page 161.

It is certain that this volume of *Structure reports* will prove invaluable, as were its predecessors, to all the many scientists in any way interested in structural studies.

AUDREY M. B. DOUGLAS

Temperature measurement in engineering, Vol. 1. By H. DEAN BAKER, Ph.D., E. A. RYDER, M.E., and N. H. BAKER, M.A. (London: Chapman and Hall Ltd.; New York: John Wiley and Sons Inc.) Pp. vi + 179. Price 30*s.*

This book is the first of two volumes, and, after an introduction dealing with the temperature scale, is confined to the industrial use of thermocouples.

A chapter on heat transfer calculations for use in the design of thermocouple sheaths is a notable feature for, as the authors point out, the engineering practice of design "based on quantitative calculation from established facts, as supplemented by judgment" has, in the field of temperature measurement, "all too frequently been curtailed to the exercise of judgment as tempered by optimism." Provided that the need for the use of judgment is not overlooked, the various graphs given in this section should be helpful.

Notes and comments

Elections to The Institute of Physics

The following elections have been made by the Board of The Institute of Physics:

Fellows: T. Alper, G. W. R. Ardley, J. F. G. Darby, E. R. Elliott, C. P. Haigh, T. R. Kaiser, W. C. Lister, Y. M. Simaika, N. Winogradoff.

Associates: C. K. Burchell, L. F. Curtis, A. Franks, D. Hatch, R. Hayes, F. M. Herbert, A. Kelly, G. H. Laycock, C. W. McQuillan, H. Mellors, R. Milligan, W. W. Moriarty, G. J. Morris, R. D. Nixon, F. H. Oakley, D. F. Perrens, J. R. Scott, E. V. Sherriff, R. F. Simpson, R. L. Spence, C. V. Taylor, F. E. Trewin, L. G. Wilson, G. T. Wright.

Nine Graduates, four Students and one Subscriber were also elected.

Prizes for original contributions published in the *Journal of Scientific Instruments*

The Board of The Institute of Physics has awarded Bowen Prizes of fifteen guineas for each of the following papers published in last year's volume of the *Journal of Scientific Instruments*.

Mr. A. G. Milnes and Mr. T. V. Vernon, of the Royal Aircraft Establishment, Farnborough, Hants, for their paper on *A saturable-core reference source for use with magnetic amplifiers* (p. 135, April 1953).

Mr. G. D. Dew of the National Physical Laboratory, Teddington, Middlesex, for his paper *On the preparation of plane diffraction grating replicas for helical rulings* (p. 229, July 1953).

Mr. G. D. Archard, A.Inst.P., of the Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks, for his paper on *Magnetic electron lens aberrations due to mechanical defects* (p. 352, October 1953).

These prizes are awarded to authors who are not more than 35 years of age and whose papers are judged to be the best in respect of originality, scientific value, practical utility to instrument makers and users, and presentation (including the standard of presentation of the manuscript). Money for the prizes is provided by the Scientific Instrument Manufacturers' Association of Great Britain Ltd., from the Bowen Trust Fund established by the late Mr. William Bowen.

Hydraulic transport of solids

We have received a copy of the *Proceedings of a Colloquium on the Hydraulic Transport of Coal* which contains the eight papers read and discussed at the National Coal Board's Colloquium in November 1952. The object of the symposium was to bring together academic and industrial research workers interested in the transport of solids down pipes by hydraulic means and to see if from their combined experiences and the mass of data available there was some hope of producing a "rational, consistent compendium which can in the future be used for intelligent design." The titles of the papers were: The hydraulic transport of solids; The fluid conveyance of solids in pipes; Experimental methods of studying the flow of solid/liquid systems; The hydraulic transport of coal and solid material in pipes; The development of a coal-feeder; The development of a lock-hopper for introducing coal into a high pressure pipe-line; The experimental coal-pumping installation at Carling, Lorraine; The plant in operation at Markham Colliery.

Copies of this 75-page booklet can be obtained gratis by *bona fide* enquirers from the National Coal Board, Scientific Department, Hobart House, Grosvenor Place, London, S.W.1.

Terminology of components of servo-mechanisms

The British Standards Institution has issued Section 5 *Components of servo-mechanisms* of its glossary of terms used in automatic controlling and regulating systems. This covers devices used for deviation measurement; amplifiers and components for amplifiers; computing devices; stabilizing and response modifying devices; servo-motors; tachometers; and miscellaneous components.

Copies can be obtained from British Standards House, London, W.1, price 5*s*. net.

Design and application of small digital computers

The fourth annual Eastern Joint Computer Conference and Exhibition on the design and application of small digital computers will be held from 8 to 10 December, 1954, at Bellevue-Stratford Hotel, Philadelphia. The conference is sponsored by the American Institute of Electrical Engineers, The Institute of Radio Engineers and the Association for Computing Machinery.

Proceedings of the conference will be published by the sponsoring societies and may be ordered from any one of them. Further information may be obtained from Eastern Joint Computer Conference, P.O. Box 7825, Philadelphia 1, Pennsylvania.

Journal of Scientific Instruments

Contents of the August issue

ORIGINAL CONTRIBUTIONS

Papers

A general-purpose electrometer. By R. M. Fry.
A precision micro-thermostat for crystal X-ray studies. By P. J. A. McKeown.
An apparatus for the determination of molecular weights by light scattering. By L. Harvey and D. Cleverdon.
A magnetic gas discharge tube oscillator. By J. M. Somerville.
A reflecting microscope for infra-red absorption measurements. By K. P. Norris.
Vignetting in visual instruments. By R. C. Lane and R. S. Longhurst.
A surface-scanning pyrometer. By R. B. Sims and J. A. Place.
A simply-constructed constant temperature light-scattering photometer and differential refractometer. By S. R. Caplan.
A strain gauge bridge for creep tests. By L. N. Clarke.

Laboratory and workshop notes

A simple photocell amplifier attachment for galvanometers. By M. W. Jervis.
Internal bridge calomel electrode. By S. Lewin.
Measurement of the damping of rapidly decaying transverse vibrations. By P. L. Kirby.
Capacitor with definite loss angles for checking bridge measurements at power frequency. By A. W. Stannett.

NOTES AND NEWS

Correspondence

Multiple film methods in X-ray powder photography. From R. J. Davis and W. E. Armstrong: J. Iball.
A note on the use of quenching circuits with Geiger–Müller counters. From R. L. Gordon.
Inexpensive X-ray detectors. From E. Leighton Yates.
Allowable ripple voltage in a photometer voltage source. From R. C. Steel.

Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London.

It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions.* (Price 2*s*. 6*d*. including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s*. each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

SPECIAL ARTICLE

The thermodynamic approach to solid structure

By VERA DANIEL, Ph.D., F.Inst.P., British Electrical Research Association, Greenford, Middlesex

A systematic treatment of the properties of materials can be approached at different levels, in terms of different concepts and variables. Solids consist of particles, and may be described either by molecular variables, referring to the properties of particles, or by molar variables, referring to a large assembly of particles as one whole. In practice we almost always measure molar variables such as weight, temperature, etc., but the physicist usually feels that a true understanding of the physics of solids can be obtained only in terms of molecular concepts and variables. However, an approach in terms of thermodynamics, which uses molar variables, is more fundamental, and more practically useful, than is sometimes realized. It is the purpose of this paper to demonstrate a few of the advantages of thermodynamic methods for the treatment of solids. Examples are given to show that a clear understanding of what is thermodynamically possible can save a good deal of misunderstanding, and that measurement of thermodynamic data in some cases makes it possible to predict properties of materials where other methods are unwieldy or impracticable.

The structure of a solid represents a solution of the problem how to satisfy the bonding forces of atoms, ions or molecules of given shapes and force fields, and, as a rule, more than one solution of this problem is possible. It is the merit of thermodynamics that it permits a quantitative assessment of the stability of possible structures. Thermodynamic methods are at their best in determining equilibrium conditions, i.e. the range of temperature, pressure, concentration, etc., within which a given structure is the most stable one. In the following, we shall briefly recapitulate the most important considerations which govern equilibria, and then give specific examples of thermodynamic methods applied to assemblies in equilibrium and out of equilibrium.

An assembly is in equilibrium if its appropriate thermodynamic potential has an extreme value. These potentials may be variously defined, but those most used are the free energies F and G which have minima in equilibrium. F applies to conditions of constant volume, and is defined as:

$$F = U - TS \tag{1}$$

where U is the total energy and S the entropy. G applies to conditions of constant pressure and is defined as:

$$G = U + pV - TS = H - TS \tag{2}$$

where H is the enthalpy or heat content. For solids and liquids at atmospheric pressure the term pV is very often negligible, so that F and G tend to be used indiscriminately. However, the permissibility of this procedure cannot be taken for granted in every case.

All the thermodynamic functions are uniquely determined for an assembly in inner equilibrium, and are accessible to measurement. For instance, the heat content H at a given temperature and pressure is defined by:

$$H_T - H_0 = \int_0^T c_p dT \tag{3}$$

where c_p is the specific heat measured at constant pressure and H_0 the heat content at absolute zero. The entropy is defined as:

$$S = \int_0^T \frac{c_p}{T} dT \tag{4}$$

the entropy of absolute zero being the same for any state of an assembly that constitutes inner equilibrium. The definitions given imply that the difference of free energy between an assembly in two states can be fully determined by calorimetric measurements, so that the difference of stability of two structures is in principle a measurable quantity. So far, data of specific heat down to very low temperatures are rather scarce, but their importance is being increasingly appreciated.[1]

The thermodynamic variables are defined for assemblies which are sufficiently homogeneous to have the same variables ascribed to all their parts. That is, they are defined for a phase, such as the gas phase, or, say, a solid of uniform composition and crystal structure. A material may consist of several phases, and thermodynamic considerations are the more involved the larger the number of phases in question. In equilibrium, the number of coexisting phases is limited by the number of independent components present. The assessment of the number of components is not always simple. For instance, pure water contains chemical species H_2O, H^+, OH^-, H_3O^+, $(H_2O)_2$, etc., and in principle each of these might be considered a component. This is, indeed, necessary where changes in the amount of chemical species, say, with temperature, are of importance. However, in general, water may be considered the only component present for considerations of phenomena such as melting, evaporation, etc.

The relationship between the number of independent components, n, the number of phases, p, and the number of degrees of freedom f (the number of variables that may be altered without causing a phase change) is given by the so-called phase rule:

$$n + 2 = p + f \tag{5}$$

This is, in effect, a thermodynamic equation, though it is not always recognized as such.

The phase rule implies that for a one-component system the maximum number of stable phases is three. In this case, the number of degrees of freedom is zero, which means that three phases can coexist only if pressure and temperature have specific values. Since any material always has a gas phase, the phase rule implies that, at the most, two condensed phases can coexist at a given temperature and pressure. Two phases, each characterized by an equation of state, cannot coexist over a temperature interval but must change into one another at a transition temperature. This temperature and the mode of phase transition can in some cases be determined by thermodynamic considerations on the basis of comparatively simple measurements.

In the following, we shall first consider some examples of thermodynamic methods applied to single component systems, before proceeding to the more involved assemblies with several components.

THERMODYNAMICS OF LONG-CHAIN CRYSTALS

A particularly simple use of thermodynamics can be made in the case of solid long-chain paraffins and paraffin derivatives.[2] It is an essential feature of these substances that they form homologous series, the members of which differ from one another by small steps. For instance, the chemical formula of any normal long-chain paraffin can be rendered as C_nH_{2n+2}, where the molecules consist of zig-zag chains of n carbon atoms with attached hydrogens. Other homologous series are derived from the normal paraffins by replacement of one or more hydrogens by some other group such as oxygen, OH or CH_3.

Long-chain substances crystallize in layer structures, the long chains being almost always parallel to one another within the layers. Within the limits set by these features, long-chain substances show a bewildering variety of structural forms. For instance, the normal paraffin $C_{24}H_{50}$ may exist in four different structures, one triclinic, two monoclinic and one hexagonal. The latter form, which occurs near the melting point, is comparatively disordered in that the long chains are able to rotate and/or twist in the solid. Disordered forms of this kind occur in numerous paraffins, and also in some derivatives, but certain homologous series, such as ketones, never show these forms.

The complicated morphology of crystalline long-chain substances follows in the last resort from the properties of the molecules, which are moderately flexible and bonded to one another by comparatively weak and not markedly directional van der Waals' forces. However, a quantitative interpretation of solid morphology in terms of these variables is extremely difficult and has been successful only with regard to a few special problems. On the other hand, simple thermodynamic considerations provide a clear picture of the whole field, and illuminate the reasons for the structural complexity.

The table shows measured thermodynamic data for normal paraffins of chain-lengths 6–25, in the solid state. The data are collected from various publications and are quoted for

n	T_M(°C)	$H_{TM}-H_0$	S_{TM}	ΔH_S	ΔH_M	ΔH_T
6	−94·3			121	35	
8	−56·5	45	0·397	130	43·2	
14	5·5	61	0·450	123	53·3	
15	10	70		133	63·0	10–20
20	38	78	0·473			
21	40				57·1	12·3
22	44·4				60·0	22·2
23	47·7				58·9	16·6
24	51·1				60·5	21·0
25	54·2	87	0·455			

selected values up to $n = 25$, which are, however, typical also for intermediate ones.[2] Energies are given in cal/g of substance (which is approximately 1/14 of the value per mole of CH_2 groups, i.e. per link of the chain). S is given in cal per °C per g. $H_{TM} - H_0$ and S_{TM} are heat content and entropy, defined by equations (3) and (4) at the melting point T_M. ΔH_S is the heat of sublimation, while ΔH_M and ΔH_T are the latent heats of melting and of the most important transition in the solid state.[2]

The table shows that the various energy and entropy terms change only quite slowly with chain-length, so that the free energies of a paraffin of given chain-length may be written down, approximately, on the basis of a few measurements. This feature can be used for a systematic treatment of the observed transitions. To this end, we may write down the free energy of a solid paraffin of given structure and of chain-length n as:

$$G_S = H_S - TS_S \tag{6}$$

and that of the same paraffin in the liquid state as:

$$G_L = H_L - TS_L \tag{7}$$

In general, one or the other of these two phases must have a lower free energy, but at the melting point T_M their free energies must be equal, so that $G_L(n) = G_S(n)$. This means that:

$$\Delta H_M = H_S - H_L = T_M(S_S - S_L) = T_M \Delta S_M \tag{8}$$

and gives a simple means for predicting melting points. Since all the terms are approximately proportional to n over fairly long ranges of chain lengths, we may put:

$$\begin{aligned} \Delta H_M &= \alpha n + \beta \\ \Delta S_M &= \gamma n + \delta \end{aligned} \tag{9}$$

where α, β, γ and δ are approximately constant for a given crystal structure. Equations (8) and (9) imply that the melting point for a homologous series approaches a limiting value for $n \to \infty$, as is indeed the case.

Equations (8) and (9) may also be used to predict the melting points of paraffin derivatives. A substituted group, such as COOH instead of CH_3, or CO instead of CH_2, adds to β and δ in equation (9) terms which are approximately constant for a given group. These may be measured by determining melting points of some compounds in which the group occurs, or, with some practice, they may be estimated from structural considerations.

Knowledge of terms β and δ may be used in two ways: in order to predict melting points of compounds so far unknown, or alternatively, to use measured melting points in order to derive structural features. For instance, secondary long-chain alcohols have high melting points, and this implies that the OH groups contribute a large term to the energy of bonding. Hence the melting point suggests the presence of hydrogen bonding.

The methods described for the prediction of melting points may be extended also to predict phase transitions in the solid. Equation (8) applies to any so-called first order transition with finite differences ΔH and ΔS at the transition, and most transitions, at any rate in pure paraffins, are of this kind. We may put:

$$\Delta H = T_c \Delta S \tag{8a}$$

where ΔH and ΔS now mean differences between two structural forms in the solid.

Equation (8a) offers a means of understanding why the transitions between ordered low temperature forms and disordered high temperature forms in paraffins occur only in some paraffins and derivatives, but not in others, and why these transitions are greatly influenced by impurities. The table shows that the transitions in question have fairly high latent heats and are not significantly dependent on chain-length and that their entropy [see equation (8a)] is such that T_c is near the melting point. This means that a small increase of the term ΔH, or a decrease of ΔS, can increase T_c to above the melting point, abolishing the disordered form.

The terms ΔH and ΔS may be influenced in two ways: by

substituted groups, in a way analogous to that discussed for melting, and by mixing. An example of the latter method is instructive because it shows that structural stabilities may, in favourable conditions, be deduced from purely thermodynamic considerations. Several long-chain ketones can be synthesized for a given chain-length n, with the ketogroups substituted at different links of the chain. Such "equivalent" ketones have very similar physical properties, and may be assumed to have approximately identical free energies. When p such ketones are mixed in equal proportions, it can be shown that equation 8(a), for the possible transition to a disordered form becomes modified to:

$$\Delta H = T_c'(\Delta S + R \log p) \tag{10}$$

where ΔH and ΔS are terms that apply for a pure ketone, while $R \log p$ is an entropy of mixing. This crude calculation is verified by experiment: pure ketones show no transitions, but when p is increased to five a hexagonal, disordered form is found. The method is sufficiently quantitative to allow the calculation of the contribution of the ketogroup to the bonding of the structure from transition point data, and the result is in reasonable agreement with other calculations of this energy.[2]

HIGHER ORDER TRANSITIONS

So far, we have considered only sharp, so-called first order, transitions, for which V, S, H, etc., are discontinuous at the transition. This is the most common case for pure substances, as well as the simplest one. However, important cases exist where transitions are more or less gradual, so-called second or higher order transitions. In the present brief review, we can only make a few points regarding these.

In general, a function $G(T)$ may be defined for any phase in equilibrium. If two phases change into one another at a first order transition, their $G(T)$ curves intersect at an angle, characterized by:

$$\left(\frac{\partial G_1}{\partial T}\right)_p - \left(\frac{\partial G_2}{\partial T}\right)_p = S_2(T_c) - S_1(T_c)$$

T_c being the transition temperature. For higher order transitions, the $G(T)$ curves of the two phases merge into one another at the transition point, i.e. they have a common tangent. Nevertheless, a definite transition temperature T_c always exists, and theories have been developed in terms of discontinuities of the differentials $\partial S/\partial T$, $\partial^2 S/\partial T^2$, etc., at the transition point. However, conditions are fairly complicated.[3] Experimentally, a higher order transition is characterized by the specific heat as a function of temperature resembling in shape a Greek lambda (Fig. 1), and such transitions are also called "lambda" transitions. Structurally such transitions represent, as a rule, some process of disordering.

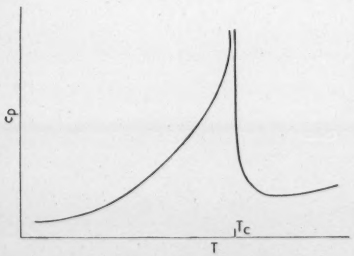


Fig. 1. Typical "lambda" transition

Without entering into a detailed discussion of the complicated subject of lambda transitions, it seems worth while to make a point which is sometimes overlooked. For such transitions c_p, the specific heat measured at constant pressure and c_v, that measured at constant volume, may be very different, so that pV terms cannot be neglected. This is intelligible from the structural point of view; a process of disordering tends to increase the volume, and an increase of volume releases the restraint of thermal vibrations which contribute to the heat content.

FERROELECTRICITY AND ANTI-FERROELECTRICITY

Ferroelectrics and anti-ferroelectrics are two classes of substances which have in common phase transitions connected with anomalies of the dielectric constant. These transitions may be either of the first or of higher orders, a gradual approach to the transition being of interest in that it may imply very high dielectric constants.

In the case of ferroelectrics, at least one of the phases in question consists of crystals which are electrically polarized. This means that even in the absence of an electric field the material consists of so-called domains, each having a polarization P per unit volume. (This is analogous to ferromagnetism, hence the curious name ferroelectrics, but the analogy is not very close. The existence of polarized domains is not in itself sufficient to imply ferroelectricity, but is common to all pyroelectric crystals. Ferroelectrics are characterized by domains, the polarization of which is easily affected by an applied field.)

The structural reasons for ferroelectricity are very complex, and different materials may owe their ferroelectric character to different reasons. However, from the thermodynamic point of view, the behaviour of all ferroelectrics can be treated in the same way,[4] by considering the free energy as a function of the polarization P.

It follows from first principles that the electric polarization of a unit volume contributes to the free energy a term which is a function of P^2. The expression for the free energy of a material contains, in general, a number of terms due to various causes. However, a theory of ferroelectricity need not be concerned with all of these. The most striking difference between two phases which change into one another at a ferroelectric transition is a change of polarization P. In the simplest case a polarized phase with finite P changes into a non-polarized phase with $P = 0$. To treat such transitions, the free energy may be expanded as a function of P^2, amalgamating all other terms in a term G_0; that is, we may put:

$$G = G_0 + \alpha P^2 + \beta P^4 + \ldots \tag{11}$$

where G_0, α, β, etc., are independent of P.

Equation (11) is a formal relationship, the value of which is greatly enhanced if data are available regarding G_0 and the coefficients α, β, etc. However, the equation is useful even in the absence of such data. For instance, it is possible to derive from it the dependence of a transition temperature between two phases, T_c, on the strength of an applied electric field, E. For a first order transition:

$$\frac{\partial T_c}{\partial E} = -\frac{P_1 - P_2}{S_1 - S_2} \tag{12}$$

where P_1, P_2 and S_1, S_2 are the polarization and entropies of the two phases respectively. The sign of the differences in question is often obvious from very simple considerations. For instance, for a change of a polarized low temperature form, phase 1, to a non-polarized high temperature form, phase 2, it is clear that $P_1 - P_2$ is positive; it is clear also that $S_1 - S_2$ is negative, since the entropy increases in going from a low temperature to a high temperature form. Hence, an external field raises the temperature of a ferroelectric transition.

Equation (12) is typical of a number of thermodynamic relationships; for instance, for any first order transition

$$\frac{\partial T_c}{\partial p} = \frac{T(V_1 - V_2)}{H_1 - H_2} \qquad (12a)$$

where p and V are pressure and volume, and H the heat content. This relationship is known as the Clausius–Clapeyron equation. Equations of this type are based on the fact that the molar variables H, S, etc., or their combinations such as G, are in equilibrium uniquely determined for a given state of an assembly. In other words, a difference such as ΔH, characteristic of two states of the assembly, is a total differential in terms of two or more variables sufficient to describe the assembly. These independent variables may be chosen in different ways, e.g. H may be expressed as $H(T, V)$ or $H(T, p)$ or $H(S, V)$, etc. By using the fact that ΔH is a total differential and by employing different combinations of independent variables, a number of relationships analogous to the Clausius–Clapeyron equation may be derived.

For an assembly in an external field (electric, gravitational, etc.), the applied field introduces a third independent variable, apart from the two variables necessary to describe an assembly in the absence of a field. This does not raise basic difficulties, and thermodynamic relationships can be derived incorporating the variables of the electric field.

The distinctive feature of the thermodynamic theory of ferroelectricity is the use of P, the polarization per unit volume of a single domain crystal (proportional to the polarization of a crystallographic unit cell) as a thermodynamic variable. Once it has been proved that this is permissible, the mathematical apparatus of thermodynamics automatically yields relationships of the type given by equation (12)—no further physical assumption is necessary, except that the structural transition envisaged is of the first order.

When it is not assumed that transitions in materials characterized by a molar variable P are of the first order, thermodynamic methods can be used to derive what kinds of parameters in equation (11) lead to first order transitions, and what kinds of parameters lead to transitions with continuous changes of the molar variables. Investigations of this kind are instructive in giving rules to the kinds of transitions possible.[5] However, they are too complicated to be discussed here.

The treatment of ferroelectricity has been extended also to the case of anti-ferroelectric materials, which have no permanent polarization P. However, it has been shown that an analogous molar variable may be used which characterizes a displacement of dipoles within a unit cell, where the resultant dipole moment is zero. Similar results may be derived as for ferroelectrics. For instance, it can be shown that, in analogy to equation (12), an applied electric field lowers the temperature of a first order anti-ferroelectric transition. Thus, the dependence of transition temperature on field strength may be used to distinguish between the ferroelectric or anti-ferroelectric character of a transition. The derivation of this practical result is based simply on the proof that P, or the anti-ferroelectric displacement, may be used as a thermodynamic variable.

The thermodynamics of ferroelectric materials shows the limitations, as well as the advantages, of the thermodynamic approach. It provides some very valuable general relationships between molar variables, but these are not sufficient to provide a full understanding of the phenomena. Beyond the relationships which the method provides, as it were, without effort, further progress can be made only by assumptions based on structural considerations. If this is done, the method loses its sovereign infallibility.

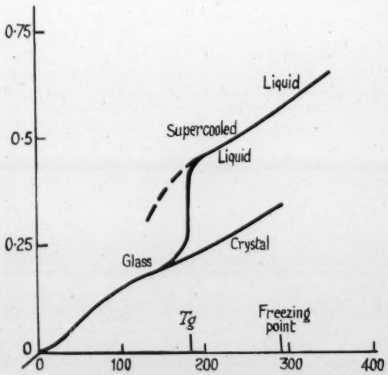
THE NATURE OF A GLASS

Methods such as those applied to long-chain crystals and ferroelectrics apply to phases in equilibrium with defined thermodynamic functions, and do not apply to materials not in equilibrium, such as glasses. However, even in those cases the thermodynamic method may be remarkably powerful. The clearest description of the nature of a glass may be given in terms of thermodynamics.[6]

Qualitatively, a glassy material is very similar to a supercooled liquid, and superficially the only difference appears to be in the magnitude of the viscosity. However, thermodynamics shows a basic difference between a glass and a supercooled liquid. Equations (4) and (8) imply a certain relationship between the latent heat of melting and the specific heat curves of the solid and a supercooled liquid, provided that both are in inner equilibrium. Thus,

$$\Delta S_M = \int_0^{T_M} \frac{C_{pL}}{T} dT - \int_0^{T_M} \frac{C_{pS}}{T} dT \qquad (13)$$

where C_{pL} and C_{pS} are the specific heats of liquid and solid respectively, and ΔS_M the measured entropy of melting. This equation is not valid for a glass. Measurement of the specific heats of glass, crystalline, solid and liquid gives the picture



[Reproduced from *Advances in Physics*[6]]

Fig. 2. Specific heat and state of aggregation

shown in Fig. 2. Instead of following a curve such as that drawn in dotted lines, which would satisfy equation (13), the specific heat drops to a low value around a temperature T_g which characterizes the change from a normal supercooled liquid to a glass. The value of $\int_0^T \frac{C_p}{T} dT$ for the glass is too low to satisfy equation (13), and this implies that a glass retains a finite entropy at absolute zero, a condition impossible for a substance in inner equilibrium.

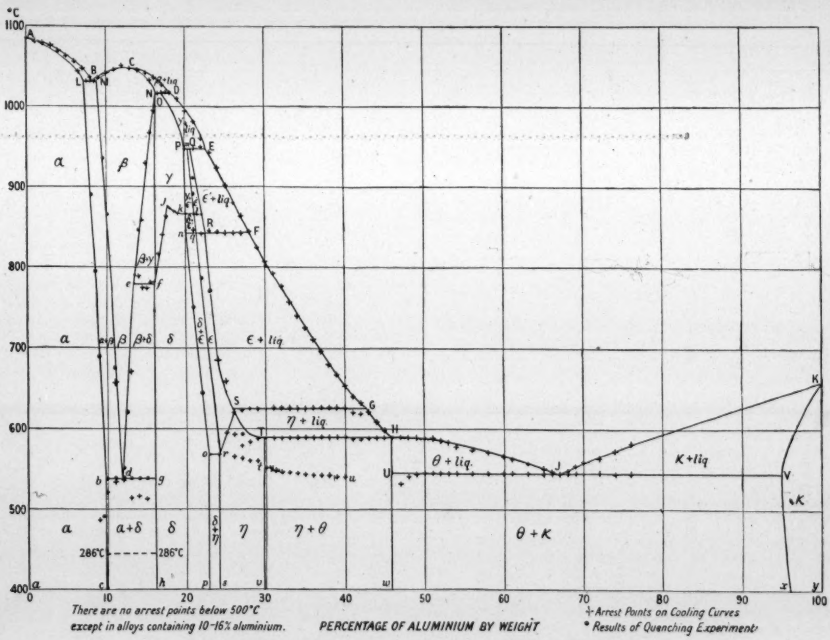
The fact that a glass is not in inner equilibrium is fundamental for any theory of the glassy state, but it has not always been appreciated. This has led to a number of misconceptions. In particular, the drop of the specific heat temperature curve at T_g has been confused with an equilibrium transition, such as those discussed above, while the two cases are entirely different. For an equilibrium transition a clearly defined state of the assembly corresponds to a given temperature and pressure. This state is independent of time, in principle. The glass, at constant temperature, must change in time, however slowly. To reach stability it must become crystalline, but, failing this, it changes so as to approach the super-cooled liquid, which is a phase with definite thermodynamic data, even though its free energy is higher than that of the crystalline solid. The state of a glass can, in principle, be described only as a function of time.

THE THERMODYNAMICS OF ALLOYS

The thermodynamic treatment of systems of several components is more involved than for one component, and the mathematics used is often of an unattractive complexity. However, the presence of several components does not introduce any fundamental difficulty, and thermodynamic methods have proved extremely useful for systems with several components. Of these, perhaps the most important are alloys, and in the following some applications of thermodynamics will be discussed, with particular reference to binary alloys.

According to the phase rule [equation (5)], a system of two components may have two condensed phases in equilibrium while retaining a degree of freedom, and thus binary alloys may be single- or two-phase. Three phases may coexist only at a singular point analogous to the melting point of a pure substance. The range of existence of the different phases as a function of temperature and concentration is usually tabulated in the form of a so-called phase diagram. Fig. 3 shows a particularly complicated example, the phase diagram of copper-aluminium alloys. Areas inscribed α, β, etc., signify single-phase regions of the respective phases, while areas such as $\alpha + \beta$ and $\beta + \gamma$ are two-phase regions.

Phase diagrams can be derived from a knowledge of the free energies of the relevant phases. For each phase, of given crystal structure but variable composition, a free energy may



[Reproduced from *Journal of the Institute of Metals*, **51**, p. 131 (1933).]

Fig. 3. Phase diagram of copper-aluminium alloys

be defined as a function of concentration per mole of material in that phase. Fig. 4 shows such functions for several phases, possible, though not necessarily stable, for an alloy system; and also shows a simple geometrical construction by which the conditions of minimum free energy may be derived. Where a common tangent, applied to two free energy curves, is lower than the curve for any pure phase, the material is two-phase in equilibrium consisting of (variable) amounts of two phases, each of fixed composition. Fig. 4 shows various interesting points. The single-phase range of a given phase need not include the minimum of its free energy curve (see phase γ). This means that the range of homogeneity of a phase need not be an indication of its ideal composition, i.e. that most favourable for reasons of bonding. Fig. 4 also shows that a strongly curved free energy curve tends to cause a narrow range of homogeneity.

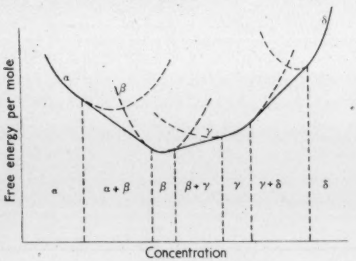


Fig. 4. Construction of phase diagram from free energy curves

The conditions of equilibrium between two phases may be defined conveniently in terms of partial differentials of the free energy. For a binary alloy with m atoms of the component A, and n atoms of component B, and two phases, 1 and 2, the condition of equilibrium is determined by the two equations:[8]

$$\frac{\partial F_1}{\partial m_1} = \frac{\partial F_2}{\partial m_2} \; ; \; \frac{\partial F_1}{\partial n_1} = \frac{\partial F_2}{\partial n_2} \qquad (14)$$

where F_i is the free energy for an alloy with $m_i + n_i$ atoms. (F, and the free energy per mole $f = F/(n+m)$, may easily be confused. The latter is used for the construction in Fig. 4.)

The partial differentials in question are related to the partial vapour pressures of A and B over the alloy. The equilibrium conditions might in fact have been stated as follows: the partial pressure of A must be the same over both phases and, at the same time, the partial pressure of B must also be the same over both phases (though, of course, the partial pressures of A and B need not be equal to one another).

Vapour pressures are, in principle, measurable quantities, and may be used to measure the partial differentials of the free energy directly. These differentials are also related to the so-called equilibrium electrode potentials, which also can be measured.[8] However, both kinds of measurements are easily vitiated by the circumstance that the surfaces which determine pressure and electrode potentials may not be representative of the interior of the alloy. Nevertheless, the possibility of such measurements should be borne in mind, since with suitable precautions they can yield valuable results.

As a rule, neither the free energies nor their partial differentials are available for alloy systems, and phase diagrams are determined experimentally. Nevertheless, thermodynamic considerations are most important because they show that the diagrams must obey certain rules. Some experimental phase diagrams have been proved false by thermodynamic argument, which excludes the possibility of certain juxtapositions of phases in equilibrium. The relevant thermodynamic rules have been derived not only for binary but also for ternary and quaternary alloys.[9]

THE THERMODYNAMICS OF A SIMPLE TYPE OF BINARY ALLOY

In the present article a more detailed consideration of the thermodynamics of alloys will be confined to one special case, where the free energy can be derived from simple theoretical assumptions. We consider an idealized binary alloy where all phases have equal crystal structure, where all thermodynamic functions are symmetrical in the components A and B, and only nearest neighbour interactions are considered.[10] With these assumptions f for a concentration $\alpha = n/(n+m)$ may be written:

$$f(\alpha) = VZ\alpha(1-\alpha) - RT\left[\alpha \log \alpha + (1-\alpha)\log(1-\alpha)\right] \qquad (15)$$

where Z is a constant and

$$V = V_{ab} - \tfrac{1}{2}(V_{aa} + V_{bb}) \qquad (15a)$$

V_{aa}, V_{bb} and V_{ab} are energies (taken positive) of pairs of atoms AA, BB and AB respectively. V is related to the heat of mixing. If it is positive, unlike atoms attract one another more than like atoms. If it is negative, the mixture tends to segregate into two phases. In the latter case, the phase diagram is characterized by a "solid solubility" gap, as shown in Fig. 5. Above a certain temperature, T_K, proportional to

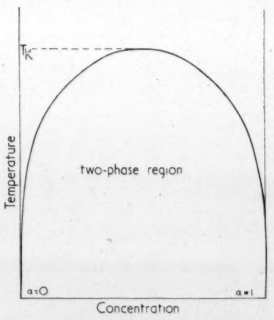


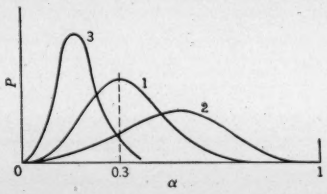
Fig. 5. Idealized "solid solubility gap"

V, the mixture is always single-phase, below this temperature it is two-phase within a range of concentration that widens with falling temperature.

The simplified model of free energy conditions characterized by equation (15) forms the basis of a thermodynamic approach to nucleation, that is, to the mechanism by which an initially single-phase solid solution breaks up into two phases. This process, which is of great importance for the hardness of alloys, is extremely complex. It involves mechanical stresses set up by the formation of nuclei of the new

phases, and is greatly dependent on texture, so that theoretical treatment in terms of molecular variables is very difficult. However, a theory in terms of thermodynamics offers the possibility of treating the process in its simplest form.

The thermodynamic treatment of nucleation starts out from the existence of fluctuations of the concentration, which exist even in a homogeneous mixture. Smoluchowski[10] shows by a thermodynamic argument on the basis of equation (15) that these fluctuations depend on V. If V is positive, i.e. if unlike atoms attract one another more than like ones, fluctuations are reduced, the more so the higher the absolute value of V. On the other hand, if V is negative, fluctuations are increased (see Fig. 6). The fluctuations in equilibrium



[Reproduced from *Phase Transformations in Solids*[10]]

Fig. 6. Probability P of finding in an alloy consisting of 30% A and 70% B, a fraction α of atoms A among the neighbours of an atom B

(1) random, (2) $V > 0$, (3) $V < 0$.

in a single-phase alloy are the higher the larger V/RT, and on approaching the boundary between the single- and two-phase states they tend towards infinity. For the two-phase state the term "fluctuation in equilibrium" has a meaning only when applied to one of the two phases, rather than to the alloy as a whole.

The theory discussed is able to provide data on microscopic texture within a solid because it employs statistical, as well as thermodynamic, concepts. Hybrid assumptions of this kind may be used for kinetic theories, such as the theory of diffusion in two-phase alloys. An idealized alloy, the free energy of which is given by equations (15, 15*a*), consists in equilibrium at a temperature T_1 below T_K, of two phases of concentrations shown in Fig. 5. When the alloy in equilibrium at a temperature T_1 is now brought to a different temperature, diffusion may either increase or decrease concentration differences: it decreases them for $T_2 > T_1$ and increases them, i.e. proceeds "uphill" against the gradient of concentration, for $T_2 < T_1$. An equation may be derived for the coefficient of diffusion for the general case on the basis of equations (15, 15*a*). This has been tested experimentally for an alloy in the centre of a solid-solubility gap where the two phases form fairly regular lamellae, and was found correct, in first approximation for diffusion both "uphill" and "downhill".[11] However, this kind of application of thermodynamics is very far removed from the absolute validity of the conditions for phase-diagrams. It stands and falls with the underlying statistical and kinetic assumptions.

GENERAL CONSIDERATIONS

Thermodynamics enters the theory of solids in a variety of ways, of which the present article can give no more than a few examples. The ultimate reason for this variety is that thermodynamics exists at two levels. Thermodynamic data, such as specific heat, vapour pressure, etc., are down to earth, measurable quantities. On the other hand, thermodynamic ideas constitute a very abstract and general framework for the understanding of nature, and thermodynamic functions may be formulated as an aid to thought, even where there is no possibility or intention of measuring them in practice.

Apart from the dualism between, as it were, abstract and concrete thermodynamics, the variety of uses of thermodynamics for solid state physics is also due to the fact that thermodynamics is often coupled with statistics, to a greater or lesser extent. The two divisions mentioned influence both the scope of the applications, and their validity. At one end of the scale, most of the phenomenological thermodynamics of long-chain crystals sets out from measured data and arrives at measurable data without the use of statistics, and without ambiguity. The theory of phase diagrams, and to some extent that of ferroelectricity, use neither statistics nor measured thermodynamic data, and in consequence they can do no more than derive the possible ways in which variables may be interrelated. The framework provided has to be filled in by statistical assumptions or by measurements. In some applications, such as the theory of nucleation, statistical assumptions can be made to overcome the essential vagueness of thermodynamics, and lead to detailed information on intricate structural features. However, the validity of the results depends on the validity of the statistics.

All the diverse methods quoted so far refer to systems in equilibrium. More variety still is introduced when we consider also the application of thermodynamics to systems not in equilibrium and to irreversible processes. However, even here it is possible to use measured data to arrive at clear cut results, as is shown by the treatment of the glassy state.

Perhaps the most significant conclusion to be drawn from the examples quoted in this article is the value of using measured thermodynamic data, such as specific heats, melting points, and, with some caution, vapour pressures. These data are often easily available, and with the help of some facility in the use of thermodynamic concepts they can be made to yield information which it may be difficult or impossible to obtain in other ways.

REFERENCES

(1) Simon, F. E. *Research*, **6**, p. 51 (1953).

(2) Daniel, V. *Advances in Phys.*, **2**, p. 450 (1953).

(3) Temperley, H. N. V. *Sci. Progr.*, **39**, p. 27 (1951).

(4) Devonshire, A. F. *Rep. Brit. Elect. Res. Assoc.*, Ref. L/T 298 (1953).

(5) Devonshire, A. F. *Phil. Mag.*, **40**, p. 1040 (1949).

(6) Davies, R. O., and Jones, J. O. *Advances in Phys.*, **2**, p. 370 (1953).

(7) Lipson, H., and Wilson, A. J. C. *J. Iron and Steel Inst.*, **2**, p. 107 (1940).

(8) Dehlinger, U. *Chemische Physik der Metalle und Legierungen* (Leipzig: Akademische Verlag gesellschaft, 1939).

(9) Wilson, A. J. C. *J. Inst. Metals*, **69**, p. 1 (1943).

(10) Smoluchowski, R. *Phase Transformations in Solids*, p. 149. (New York: J. Wiley and Sons Inc.; London: Chapman and Hall Ltd., 1951).

(11) Daniel, V. *Proc. Roy. Soc. A*, **192**, p. 575 (1948).

Methods for studying the thermal resistances of sprayed and electro-deposited metal coatings

By R. W. Powell, D.Sc., Ph.D., F.Inst.P., and M. J. Hickman, National Physical Laboratory, Teddington, Middlesex

[*Paper received* 9 *February*, 1954]

Two methods are described for investigating the additional thermal resistance introduced when the plane end of a metal bar is coated with sprayed steel or an electro-deposited metal. One is applicable to unground surfaces and involves the use of a thermopile for surface temperature measurement. The other can only be used on flat surfaces and is a modification of the divided bar method. When applied to coatings a few millimetres thick, this method enables determinations to be made both of the thermal resistance of the deposited metal and of its bond to the underlying metal.

The thermal conductivity of sprayed steel was about one-seventh of that of a similar steel in the normal form.

Some time ago the writers had occasion to investigate the additional thermal resistances introduced when certain types of coatings were applied to the surfaces of metals. The particular problem arose in 1941 and related to aluminium alloys coated with sprayed carbon steel, information being required for both the interfacial resistance between the two metals and the additional resistance due to a known thickness of the sprayed steel. For the purposes of the tests the sprayed coatings were deposited on the plane ends of $\frac{7}{8}$ in. diameter bars of the aluminium alloy. Two methods were used and are described in the account which follows.

The first, which involved the measurement of surface temperatures by a calibrated thermopile, was primarily intended for tests on coatings in the "as deposited" condition. The second was a modification of the divided bar method of thermal conductivity measurement, which could only be used after the surfaces had been ground and lapped to optical flatness. As the machining involved in the preparation of the test sample was liable to impair the bond with the underlying metal, the second method was mainly used in experiments designed to determine the thermal conductivity of the coating material. This method has since been used when determining the thermal conductivity of electro-deposited metals, and results obtained for chromium and nickel deposits are included.

THE THERMOPILE METHOD

Figs. 1 and 2 illustrate the experimental arrangement and the sequence of operations used with the thermopile method. The test bar, with thermocouples composed of 40 s.w.g. nickel-chrome and eureka wires attached at points along its length and an electric heating coil slipped over the uncoated end, is mounted as shown in an insulated container with the coated surface facing the sensitive elements of a thermopile. The thermopile is housed in a water-cooled enclosure, and diaphragms and distances are arranged so that the thermopile receives radiation from a circular area having a diameter of about 0·9 of the full diameter of the bar. It is convenient to mount the units on supports which can slide on an optical bench. To calibrate the thermopile for the particular conditions of each test the bar is heated to a steady uniform temperature, as shown in Fig. 2(*a*), when the temperature of the coated surface corresponds to the temperature indicated by the thermocouples. The front heater is then removed, the bar slid into its standard position [Fig. 2(*b*)] and the e.m.f. of the thermopile noted on a potentiometer as speedily as possible. When necessary, a correction is applied to allow for the cooling of the surface. Immediately before and after each observation a zero reading is taken by sighting the thermopile on a blackened water-

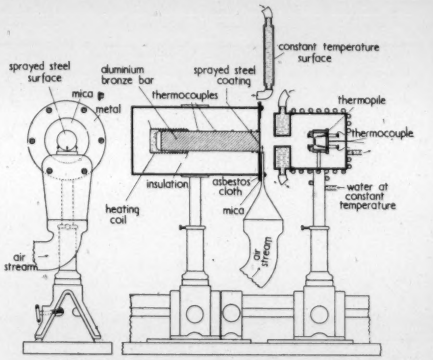


Fig. 1. Method for testing unground surfaces

cooled disk maintained at approximately the same constant temperature as the thermopile. By repeating the experiment for a series of steady temperatures a curve can be obtained relating the thermopile reading with the temperature of the surface of the coating. A forced air-stream is then directed across the face of the bar, as shown in Figs. 1 and 2(*c*), and a gradient of temperature established in the bar. When steady conditions are reached, the bar is again moved to its standard position (Fig. 1), and the thermopile reading observed. This reading gives the temperature of the external surface of the coating on the end of the bar, whilst an extrapolation of the curve drawn through the temperatures indicated by the thermocouples pegged into the bar gives the temperature of the interface between the base-metal and the coating. Hence the total temperature drop across the coating and its bond to the base-metal can be determined. The quantity of heat

flowing can be derived from the observed temperature gradient in the bar and a knowledge of the thermal conductivity of the metal constituting the bar.

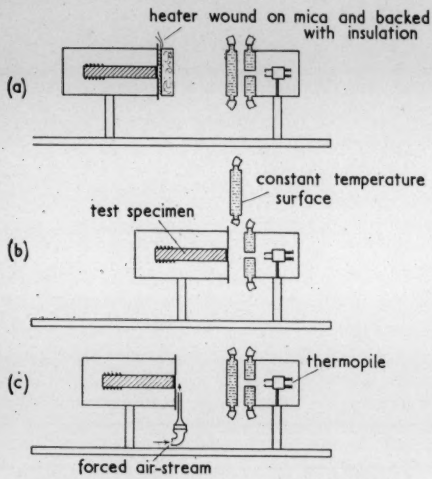


Fig. 2. Disposition of apparatus for thermopile method

(*a*) Steadying for calibration experiment; (*b*) In position for observation; (*c*) Steadying for gradient experiment

Preliminary measurements were made with an uncoated bar of graphite maintained at temperatures of the order of 200° C. These confirmed that surface temperatures, as determined by the thermopile method, agreed to within about 0·5° C with those obtained by extrapolation of the thermocouple readings.

THE MODIFIED DIVIDED BAR METHOD

The apparatus used for the divided bar method is shown in Fig. 3. The coated bar is mounted vertically on a bar of hardened steel of the same diameter, the lapped surface of the coating being wrung on to a lapped surface of the steel bar. The bars are lagged and a gradient of temperature is established from top to bottom. The temperature discontinuity introduced by the presence of the coating is derived from a plot of the temperatures indicated by the thermocouples pegged into the bars against their positions along the length of the bars.

So far the procedure is as for a normal divided bar apparatus except that no guard tube is used.[1] The modification consists in the removal of sufficient lagging to expose the coating and short lengths of the adjacent bars. A thermocouple consisting of 40 s.w.g. nickel-chrome and eureka wires butt-welded together[2] is looped around the bars. This is used to measure the bar temperatures at a number of positions, which can be noted by a travelling microscope. The presence of an oxide film on the wires prevents the thermojunction from being partially shorted out by the bar itself. Thermal contact between the thermocouple and the bar is improved by the application of a film of oil, and a mica shield prevents the readings of the thermocouple from being seriously affected by draughts and convection currents. A potentiometer reading in microvolts is used for all thermocouple observations.

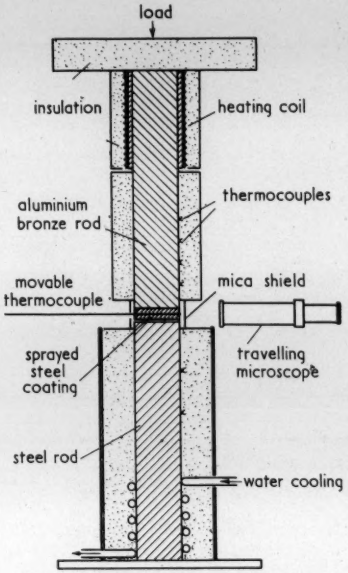


Fig. 3. Apparatus for divided bar method

Fig. 4 contains the results of a typical experiment, conducted in this way, on a coating of sprayed steel of sufficient thickness for its thermal resistance to be differentiated from the contact resistances. The results in Fig. 4 show the degree of agreement obtained between the readings of the movable and fixed thermocouples. It will be seen that the use of the movable thermocouple enables measurements to be made, not only of the temperature gradient in the coating,

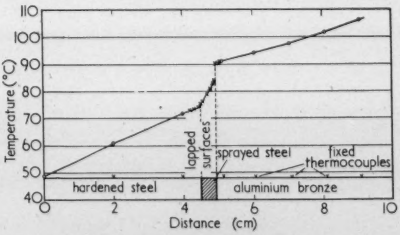


Fig. 4. Divided bar method applied to bar No. 17 (thick sprayed steel coating)

○, fixed thermocouples readings; ×, movable surface thermocouple reading

but also of the temperature discontinuities at its boundaries. In the example shown there is no appreciable discontinuity at the wrung surfaces, but there is evidence of a considerable thermal resistance between the coating and the underlying metal.

EXPERIMENTAL RESULTS

(*a*) *For thick coatings of sprayed steel.* The modified divided bar method has been applied to two thick samples of 0·75% carbon steel when sprayed on to two different metal bars. The results are set out in Table 1.

The mean values obtained for the thermal conductivities of the two samples of sprayed steel were $0{\cdot}07_5$ and $0{\cdot}06_4$ j.cm/cm² sec °C. These may be compared with a value of 0·48 for the thermal conductivity of a normal 0·8% annealed carbon steel at the same temperature.[3] The density of a similar sprayed steel sample prepared at the same time was 6·65 g/cm³, which is about 15% less than the density of the solid steel with which comparison has been made.

A microscopical examination carried out in the Metallurgy Division of this Laboratory showed the sprayed layer to consist of steel, thin oxide films and cavities. It is presumably the presence of oxides and cavities which causes the low thermal conductivity. In the light of a subsequent investigation,[4] into the conductivity of porous bronzes it seems that the low conductivity of the sprayed steel is more likely to be due to the presence of the oxide than to the porosity.

It will be observed that for these samples there is a large and variable contact resistance between the coating and the underlying metal and that impregnation with oil caused a reduction in this resistance. It seems probable that the contact had been impaired during the machining and grinding of the thick specimen. Similar measurements made on bars coated to thicknesses of 0·02 to 0·04 cm gave contact resistances which were equivalent to 0·2 to 0·3 cm of aluminium bronze and confirmed that much larger contact resistances were associated with the thick coatings.

Table 1. *Results for thick coatings of 0·75% carbon sprayed steel*

Description of sample	*Mean temperature of sample (° C)*	*Position of measurement (angle round circumference)*	*Thermal conductivity of sprayed steel (j. cm/cm² sec °C)*	*Contact resistance (Equivalent thickness in cm of aluminium bronze of conductivity 0·73 j. cm/cm² sec °C)*
Bar No. 17, 0.43 cm coating on RR aluminium bronze	80	0	0·085	2·3
		50	0·08	2·0
		110	0·075	1·6
		160	0·07	1·3
		290	0·07	1·7
Bar No. 17, 0·43 cm coating on RR aluminium bronze	170	0	0·08	2·4
Same sample after oil immersion	170	0	0·085	1·7
Bar No. 45, 0·53 cm coating on aluminium alloy RR 50-	80	0	0·065	0·65
		180	0·062	0·6

(*b*) *For thin coatings of sprayed steel.* The results of tests made by both methods on three bars coated thinly with the sprayed steel are given in Table 2.

The coating thicknesses given in the third column have been derived from the differences in the length of the bar before and after spraying. For the as received condition the coating was not uniform in thickness. After grinding the coatings of bars Nos. 1 and 2 were visibly separated from the bars at points around the edges. This would explain the higher values obtained for these two bars by the divided bar method and the dependence of the thermal resistance on loading. The thermopile would not receive radiation from the edge of the disk where these effects were greatest. With bar No. 3, where no gaps could be observed between the

Table 2. *Results for thin coatings of sprayed steel on aluminium bronze*

Bar no.	*Condition*	*Coating thickness (cm)*	*Thermal resistances expressed as equivalent thickness of aluminium bronze at 210° C (cm)*					
			Total effect of coating				*Due to sprayed steel*	*Due to interfacial contact*
			Thermopile method	*Divided-bar method*		*Mean (weighted)*		
				Unloaded	*Loaded*			
3	As received	0·033	0·45	—	—	0·45	0·34	0·11
3	Ground	0·02	0·48	0·43	0·41	0·45	0·21	0·24
2	Ground	0·04	0·8	1·6*	0·6	0·7	0·42	0·28
1	Ground	0·024	0·43	1·4*	1·0*	0·43	0·25	0·18

* High values due to partial separation of coating from bar.

coating and the base-metal, there is less dependence on loading, and better agreement between the two methods. The eighth column gives the thermal resistance due to the observed thickness of sprayed steel assuming it to have a thermal conductivity of 0·07 j.cm/cm² sec °C, the mean of the values obtained for the two thick samples. The last column shows the thermal resistance between the coating and the base-metal, being the differences between the data given in the two preceding columns. There is again indication that the grinding has impaired the contact. It is possible that this might have been avoided had the special grinding techniques now available been used.

(*c*) *For electro-deposited metals.* The divided bar method, when modified as described above, has also been used for coatings of electro-deposited metals. The first such coating was one of nickel-tin alloy on a bar of aluminium bronze which, after lapping, had a thickness of only about 0·08 cm. Measurements made with the coating at a temperature of about 230° C, showed the thermal contact between the electro-deposited metal and the base-metal to be good and indicated the thermal conductivity of the nickel-tin deposit to be of the order of 0·7 j.cm/cm²sec °C, but the deposit was too thin for more than a very rough value to be obtained.

In connexion with another investigation, the method has since been used for thermal conductivity determinations on insoluble anode process (I.A.P.) nickel and hard chromium when electro-deposited to thicknesses of 0·39 and 0·28 cm on the lapped ends of steel bars. The temperature explorations were made in three positions, *A*, *B* and *C*, which were equally spaced around the bars. No sudden drop in temperature was observed at the contact between the plating and the underlying metal, which again confirmed that the electro-deposited metals were in good thermal contact with the base-metal. The wrung contacts between the pairs of bars were generally good, but were not always uniform and they showed a tendency to deteriorate as the temperature was raised. Such variations affected the temperature gradients near the joint and limited the accuracy of the thermal conductivity determination. Table 3 contains the results obtained for these two bars.

Table 3. *Thermal conductivity of electro-deposited metals*

Metal	*Temperature* (° C)	*Thermal conductivity* (*j. cm/cm² sec °C*)		
		Position	*Value*	*Mean value*
Hard chromium	35	*A*	0·36	
		B	0·38	0·36
		C	0·34	
Hard chromium	43	*A*	0·40	
		B	0·36	0·37
		C	0·34	
Hard chromium	73	*A*	0·36	
		B	0·32	0·32
		C	0·29	
I.A.P. nickel	35	*A*	0·75	
		B	0·75	0·74
		C	0·71	
I.A.P. nickel	58	*A*	0·75	
		B	0·67	0·71
		C	0·71	
I.A.P. nickel	120	*A*	0·59	
		B	0·67	0·63
		C	0·63	

On attempting to extend the work to higher temperatures the bars tended to separate and a temperature discontinuity developed at the lapped surfaces. In the case of the hard chromium it was subsequently found that a transverse crack had developed. The coating had been heated to about 130° C when this occurred.

The thermal conductivity values obtained for electrolytic nickel are about 20% lower than were previously obtained[5] for another sample of electrolytic nickel, but the fall in conductivity with increase in temperature is similar.

There is very little published data for the thermal conductivity of chromium. Söchtig[6] has obtained values of the order of 0·27 at temperatures comparable with those of the present investigation. Work now in hand on another electrolytic sample is showing a marked dependence of thermal conductivity on subsequent heat treatment. Söchtig's sample had been heated to 1000° C for 30 min.

OTHER APPLICATIONS

The present form of the divided bar method might usefully be applied to other heat transfer problems such as the measurement of the effective thermal resistance between metal surfaces having various degrees of roughness. The effects of oxide and similar surface films might also be studied.

ACKNOWLEDGEMENTS

The work described above was carried out in the Physics Division of the National Physical Laboratory. The spray steel coated bars were supplied by Mr. J. F. Alcock of Ricardo and Co. Engineers (1927) Ltd., the aluminium bronze bars were coated by Metal Sprayers Ltd. and the bar of RR 50 aluminium alloy by Metallisation Ltd. This work was carried out in 1941 and the results apply to spray deposits formed at that time. The paper is published by permission of the Director of the Laboratory.

REFERENCES

(1) Powell, R. W., and Griffiths, E. *Proc. Roy. Soc., A*, **163**, p. 189 (1937).

(2) Hickson, V. M. *J. Sci. Instrum.*, **17**, p. 182 (1940).

(3) Powell, R. W., and Hickman, M. J. *Special Report* No. 24, p. 242 (London: Iron and Steel Institute, 1939).

(4) Grootenhuis, P., Powell, R. W., and Tye, R. P. *Proc. Phys. Soc.* [*London*] *B*, **65**, p. 502 (1952).

(5) Griffiths, E., Powell, R. W., and Hickman, M. J. *J. Inst. Fuel*, **15**, p. 107 (1942).

(6) Söchtig, H. *Ann. Phys.* [*Leipzig*], **38**, p. 97 (1940).

The capacity and field of a split cylindrical condenser, using the method of inversion

By H. J. Peake, M.A., M.Sc., and N. Davy, D.Sc., University of Nottingham

[*Paper first received* 24 *November*, 1953, *and in final form* 11 *January*, 1954]

The complex potential of a split cylindrical condenser is obtained by inversion of a known, simpler case. Expressions are obtained for the value of the electrostatic field at points on the axes of symmetry, the surface density of charge on a conductor and the capacity of the condenser. The expressions obtained by Adams, using another method, are deduced as one of three special cases for which tables and graphs are provided. The results should prove of value in the design of electrode systems for various purposes.

NOTATION

The co-ordinates of any point in the plane before and after inversion are denoted by $z = x + iy$, and z_1 respectively, $\bar{z}_1$ is the complex conjugate of z_1. The complex potential is $w = u + iv$, where u is the potential and v the stream function. The radius of inversion is a.

The usual notation is used to denote the elliptic functions of modulus k, namely sn w, but when the modulus is k', where $k'^2 = 1 - k^2$, the function is written as sn(w, k'). $K = \int_0^{\pi/2} d\phi/(1 - k^2 \sin^2 \phi)^{\frac{1}{2}}$, and K' is the same function of k' as K is of k.

THE TRANSFORMATION

The transformation

$$z = b \,\mathrm{sn}(Kw/U) \tag{1}$$

gives the complex potential at any point in the z plane for the case of two infinite plane conductors AB, CD, of finite widths $b(1 - k)/k$ placed with their sides parallel to each other and raised to equal but opposite potentials of magnitude U. The capacity of the system is $K'/4\pi K$ (Thomson) (Fig. 1).

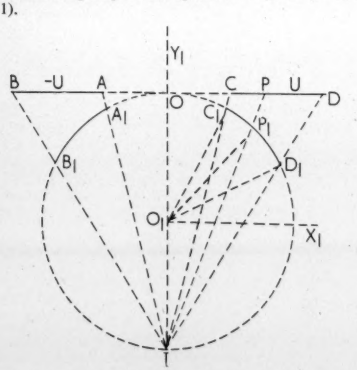


Fig.1. The geometrical construction of the inverse system

THE INVERSION

Geometrical argument. On inversion about a point I, on the axis of symmetry distant a, from AD, two equal arcs of the same circle, centre O_1 and radius $a/2$ $(= R)$, are obtained. The arcs OA_1, OB_1, OP_1 (where P_1 is any point on either arc), subtend angles $\pi/2 - \alpha$, $\pi/2 - \beta$, $\pi/2 - \theta$, at O_1 respectively. It can be shown that:

$$\begin{aligned} b^2/a^2 &= (1 - \sin\alpha)/(1 + \sin\alpha) = \tan^2(\pi/4 - \alpha/2) \\ b^2/a^2k^2 &= (1 - \sin\beta)/(1 + \sin\beta) \\ x^2/a^2 &= (1 - \sin\theta)/(1 + \sin\theta) \end{aligned} \tag{2}$$

From these equations we may determine b^2/a^2 and k^2 for given values of α and β. Indeed,

$$k = \tan(\pi/4 - \alpha/2)\cot(\pi/4 - \beta/2) \tag{3}$$

Algebraic transformation. O is taken to be the origin in the z plane and O_1 that in the z_1 plane. By definition of inversion, $IP \,.\, IP_1 = a^2$, so that

$$(z + ai) \,.\, (\bar{z}_1 - \tfrac{1}{2}ai) = a^2 \tag{4}$$

From equations (1) and (4)

$$b \,\mathrm{sn}(Kw_1/U) = -\,ai(\bar{z}_1 + ia/2)/(\bar{z}_1 - ia/2) \tag{5}$$

If $z_1 = r(\cos\theta + i\sin\theta)$, we obtain after some reduction:

$$\begin{aligned} b \,\mathrm{sn}(Kw_1/U) = 2R[2rR\cos\theta + i(R^2 - r^2)]/ \\ (R^2 + r^2 + 2Rr\sin\theta) \end{aligned} \tag{6}$$

$$dz/dz_1 = -\,a^2/(\bar{z}_1 - ia/2)^2 = -\,(z + ai)^2/a^2 \tag{7}$$

Electrical (or magnetic) *argument.* Since inversion is a particular conformal transformation, A_1B_1 and C_1D_1 may be considered conductors raised to potentials $\pm U$. Equations (5) and (6) thus give the complex potential w_1 in terms of the co-ordinate z_1 at any point in a plane normal to the axis of a split cylinder of radius R.

ELECTRICAL INTENSITY

At any point in the plane the magnitude of the electrical intensity E (units are volts per cm if U is in volts) is given by

$$E = \frac{1}{4\pi}\left|\frac{dw_1}{dz_1}\right| = \frac{1}{4\pi}\left|\frac{dw}{dz}\cdot\frac{dz}{dz_1}\right|$$

Since $(d/dw)(\mathrm{sn}\, w) = \mathrm{cn}\, w\, \mathrm{dn}\, w$, and $dz/dz_1 = -\,a^2/(\bar{z}_1 - \tfrac{1}{2}ai)^2$ from equation (7) we have:

$$E = (a^2U/4\pi bK)\big|\mathrm{cn}(Kw/U)\,\mathrm{dn}(Kw/U)(\bar{z}_1 - \tfrac{1}{2}ai)^2\big|$$

Now, $b \,\mathrm{sn}(Kw/U) = z$, so that $b^2\,\mathrm{cn}^2(Kw/U) = b^2 - z^2$, and $b^2\,\mathrm{dn}^2(Kw/U) = b^2 - k^2z^2$.

Whence;

$$E = (bU/4\pi a^2K)\big|(z + ai)^2/[(b^2 - z^2)(b^2 - k^2z^2)]^{\frac{1}{2}}\big| \tag{8}$$

The magnitude of the electrical intensity is of particular interest at points on the axis of symmetry OY_1 and along the diameter OX_1.

E for points on the *axis of symmetry*.

In this case, $z = iy$ so that $z_1 = R(a - y)i/(a + y)$. Write "t" for y_1/R. Then $y/a = (1 - t)/(1 + t)$ and equation (8) becomes

$$E = (bU/4\pi a^2 K)\{(a + y)^2/[(b^2 + y^2)(b^2 + k^2 y^2)]^{\frac{1}{2}}\}$$

On elimination of (b/ak), (b/a) by means of relation (2) and replacing y by the appropriate expression in t, we obtain the expression:

$$E = (U/4\pi KR)/[(1 + \sin\alpha)(1 - \sin\beta)/(t^2 - 2t\sin\alpha + 1)(t^2 - 2t\sin\beta + 1)] \quad (9)$$

where $t = y_1/R$.

On *the diameter*: $y_1 = 0$ and if we replace x_1 by Rt, we obtain for z:

$$2R(1 - it)/(t - i)$$

Substituting this expression for z in equation (8) and eliminating a, b and k as above, we obtain, after some reduction, that

$$E = \frac{(U/4\pi KR)\sqrt{[(1 + \sin\alpha)(1 - \sin\beta)]}}{[(t^4 - 2t^2\cos 2\alpha + 1)(t^4 - 2t^2\cos 2\beta + 1)]^{\frac{1}{4}}} \quad (10)$$

The field at the centre of the circle is obtained by putting $t = 0$ in either equation (9) or (10). The magnitude is therefore:

$$(U/4\pi KR)\sqrt{[(1 + \sin\alpha)(1 - \sin\beta)]} \quad (11)$$

SURFACE DENSITY OF CHARGE

At points on a conductor the quantity $\frac{1}{4\pi}\left|\frac{dw_1}{dz_1}\right|$ is the magnitude of the surface density of charge σ at the point under consideration. Since $z = x$ at such points ($b \leqslant |x| \leqslant b/k$) we have from equation (8) that

$$\sigma = (bU/4\pi a^2 K)(a^2 + x^2)/\sqrt{[(x^2 - b^2)(b^2 - k^2 x^2)]} \quad (12)$$

On elimination of b, k and x by means of relations (2) we obtain the expression:

$$\sigma = (U/8\pi KR)\sqrt{[(1 + \sin\alpha)(1 - \sin\beta)/(\sin\alpha - \sin\theta)(\sin\theta - \sin\beta)]} \quad (13)$$

CASES OF ESPECIAL INTEREST

We now procede to discuss three cases, which are of particular interest.

Case 1: *Symmetrical case* (Fig. 2).

In this case $\beta = -\alpha$, so that from relations (2) we have that

$$k = (b/a)^2 = \tan^2(\pi/4 - \alpha/2) \quad (14)$$

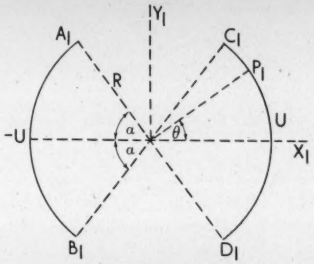


Fig. 2. Symmetrical case

From equations (9), (10) and (13) we quickly derive the following.

Electrical intensity E. At points on O_1Y_1 $\quad t = y_1/R$

$$E = (U/4\pi KR)(1 + \sin\alpha)/\sqrt{(t^4 + 2t^2\cos 2\alpha + 1)} \quad (15)$$

At points on O_1X_1 $\quad t = x_1/R$

$$E = (U/4\pi KR)(1 + \sin\alpha)/\sqrt{(t^4 - 2t^2\cos 2\alpha + 1)} \quad (16)$$

At the centre: $(U/4\pi KR)(1 + \sin\alpha)$.

Surface density of charge σ.

$$\sigma = (U/8\pi KR)(1 + \sin\alpha)/\sqrt{(\sin^2\alpha - \sin^2\theta)} \quad (17)$$

This result was obtained by Adams, using another method.

Table 1. *Values of* $(8\pi R/U)E$: *symmetrical case. Obtained from equations* (15) *and* (16)
$t = y_1/R$ on O_1Y_1; x_1/R on O_1X_1

	α \ $\pm t$	0·0	0·2	0·4	0·6	0·8	0·9	1·0	1·2	1·4	1·6	1·8	2·0	3·0	5·0
Points on 0_1Y_1	$\frac{\pi}{2}$	2·55	2·65	3·03	3·98	7·07	13·4	∞	5·79	2·65	1·63	1·14	0·85	0·32	0·11
	$\frac{\pi}{3}$	2·35	2·39	2·52	2·67	2·67	2·55	2·34	1·83	1·38	1·05	0·82	0·65	0·27	0·09
	$\frac{\pi}{4}$	2·16	2·16	2·13	2·03	1·82	1·68	1·53	1·23	0·98	0·79	0·64	0·52	0·24	0·09
	$\frac{\pi}{6}$	1·86	1·82	1·70	1·52	1·30	1·18	1·07	0·87	0·71	0·58	0·48	0·41	0·20	0·07
Points on 0_1X_1	$\frac{\pi}{2}$	2·55	2·45	2·20	1·87	1·55	1·41	1·27	1·04	0·86	0·72	0·60	0·51	0·26	0·10
	$\frac{\pi}{3}$	2·35	2·30	2·15	1·92	1·64	1·49	1·35	1·10	0·90	0·74	0·61	0·51	0·25	0·09
	$\frac{\pi}{4}$	2·16	2·16	2·13	2·03	1·82	1·68	1·53	1·23	0·98	0·79	0·64	0·52	0·24	0·09
	$\frac{\pi}{6}$	1·86	1·89	2·00	2·12	2·12	2·02	1·86	1·45	1·09	0·83	0·65	0·52	0·22	0·08

Note: The number given is the value of $(8\pi R/U)E$ for that value of t and α.

Table 2. *Values of* $(8\pi R/U)\sigma$: *symmetrical case. Obtained from equation* (17)

α \ $\pm\frac{\theta}{\alpha}$	0·000	0·250	0·500	0·600	0·700	0·750	0·800	0·850	0·900	0·920	0·940	0·960	0·980	1·000
$\frac{\pi}{2}$	1·27	1·38	1·80	2·17	2·81	3·33	4·12	5·46	8·14	10·2	13·5	20·3	40·5	∞
$\frac{\pi}{3}$	1·35	1·42	1·66	1·84	2·13	2·35	2·64	3·07	3·80	4·26	4·94	6·09	8·66	∞
$\frac{\pi}{4}$	1·53	1·59	1·81	1·99	2·26	2·47	2·74	3·16	3·86	4·31	4·84	6·08	8·60	∞
$\frac{\pi}{6}$	1·86	1·92	2·17	2·36	2·66	2·88	3·19	3·65	4·43	4·94	5·68	6·94	9·78	∞

Tables 1 and 2 give the values of $(8\pi R/U)E$ and $(8\pi R/U)\sigma$ respectively, for selected values of α.

Case 2: B_1D_1 *a diameter.*

In this case $\beta = 0$, so that

$$k^2 = b^2/a^2 = \tan^2(\pi/4 - \alpha/2) \qquad (18)$$

and equations (9), (10) and (13) quickly reduce to the following.

Electrical intensity. Points on $O_1Y_1 \quad t = y_1/R$.

$$E = (U/4\pi KR)\sqrt{[(1 + \sin\alpha)/(t^2 - 2t\sin\alpha + 1)(t^2 + 1)]} \qquad (19)$$

Points on $O_1X_1 \quad t = x_1/R$.

$$E = (U/4\pi KR)\sqrt{[(1 + \sin\alpha)/(t^2 \sim 1)]/\sqrt[4]{(t^4 - 2t^2\cos 2\alpha + 1)}} \qquad (20)$$

At the centre: $(U/4\pi KR)\sqrt{(1 + \sin\alpha)}$.

Table 3. *Values of* $(8\pi R/U)E$ *at points on* O_1Y_1 *when* B_1D_1 *is a diameter. Obtained from equation* (19)

α \ $\lvert t\rvert$	0·0	0·2	0·4	0·6	0·8	0·9	1·0	1·2	1·4	1·6	1·8	2·0	3·0	5·0
							Positive values of t							
$\frac{\pi}{2}$	1·80	2·21	2·79	3·86	7·03	13·4	∞	5·77	2·62	1·59	1·09	0·81	0·29	0·09
$\frac{\pi}{3}$	1·71	2·01	2·32	2·59	2·64	2·53	2·33	1·82	1·36	1·02	0·78	0·62	0·25	0·08
$\frac{\pi}{4}$	1·59	1·79	1·91	1·90	1·74	1·61	1·47	1·18	0·93	0·74	0·58	0·48	0·21	0·07
$\frac{\pi}{6}$	1·41	1·51	1·51	1·39	1·21	1·10	1·00	0·81	0·66	0·54	0·44	0·37	0·17	0·06
							Negative values of t							
$\frac{\pi}{2}$	1·80	1·47	1·19	0·97	0·78	0·70	0·64	0·52	0·40	0·37	0·31	0·27	0·14	0·05
$\frac{\pi}{3}$	1·71	1·42	1·16	0·95	0·77	0·69	0·63	0·51	0·43	0·36	0·31	0·26	0·14	0·06
$\frac{\pi}{4}$	1·59	1·35	1·12	0·92	0·75	0·67	0·61	0·50	0·42	0·35	0·30	0·25	0·13	0·05
$\frac{\pi}{6}$	1·41	1·25	1·05	0·87	0·71	0·64	0·58	0·47	0·39	0·33	0·28	0·24	0·12	0·05

Table 4. *Values of* $(8\pi R/U)E$ *at points on* O_1X_1 *when* B_1D_1 *is a diameter. Obtained from equation* (20)

α \ $\pm t$	0·0	0·2	0·4	0·6	0·8	0·9	0·94	0·98	1·00	1·2	1·4	1·6	1·8	2·0	3·0
$\frac{\pi}{2}$	1·80	1·80	1·82	1·93	2·34	3·07	3·85	6·46	∞	1·74	1·07	0·76	0·58	0·47	0·20
$\frac{\pi}{3}$	1·71	1·73	1·78	1·93	2·29	3·13	3·92	6·59	∞	1·77	1·08	0·77	0·58	0·46	0·20
$\frac{\pi}{4}$	1·59	1·62	1·72	1·93	2·43	3·21	4·03	6·78	∞	1·81	1·08	0·77	0·58	0·45	0·19
$\frac{\pi}{6}$	1·41	1·45	1·60	1·88	2·51	3·37	4·25	7·15	∞	1·89	1·11	0·76	0·56	0·43	0·17

Table 5. *Values of* $(8\pi R/U)\sigma$ *when* B_1D_1 *is a diameter. Obtained from equation* (21)

α \ θ/α	0·0	0·05	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80	0·90	0·92	0·94	0·96	0·98	1·0
$\frac{\pi}{6}$	∞	6·33	4·62	3·47	3·04	2·86	2·83	2·90	3·13	3·63	4·89	5·41	6·19	7·53	10·5	∞
$\frac{\pi}{4}$	∞	4·90	3·57	2·70	2·39	2·26	2·25	2·34	2·56	3·0	4·10	4·56	5·24	6·38	8·99	∞
$\frac{\pi}{3}$	∞	4·14	3·03	2·36	2·06	1·98	2·00	2·11	2·35	2·83	3·98	4·45	5·14	6·31	8·97	∞
$\frac{\pi}{2}$	∞	3·35	2·48	1·95	1·81	1·83	1·98	2·29	2·89	4·17	8·17	10·2	13·6	20·2	40·3	∞

Surface density of charge.

$$\sigma = (U/8\pi KR)\sqrt{[(1 + \sin\alpha)/\sin\theta(\sin\alpha - \sin\theta)]} \quad (21)$$

This is a minimum when $\sin\theta(\sin\alpha - \sin\theta)$ is a maximum, i.e. when $\sin\theta = \frac{1}{2}\sin\alpha$. The minimum value of σ is thus $2(U/8\pi KR)\sqrt{(1 + \sin\alpha)/\sin\alpha}$.

Tables 3, 4 and 5 give the relevant data in the case under consideration.

Case 3: $\beta = -\pi/2$ (Fig. 3).

It is convenient to replace α by γ, where $\gamma = \alpha + \pi/2$, Θ by ϕ, where $\phi = \pi/2 + \theta$, in formulae (9), (10) and (13). Since $k = 0$, when $\beta = -\pi/2$, K is then $\pi/2$ and the following results obtain.

Electrical intensity E. At points on O_1Y_1 $\quad t = y_1/R$.

$$E = (U/\pi^2 R)\sin(\gamma/2)/(t + 1)\sqrt{(t^2 + 2t\cos\gamma + 1)} \quad (22)$$

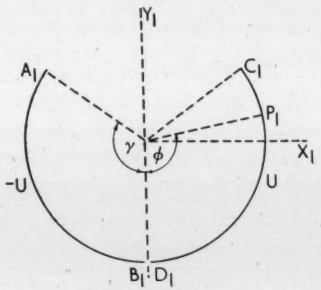


Fig. 3. Third case

Table 6. *Values of* $(8\pi R/U)E$ *for points on* O_1Y_1; $t = y_1/R$. *Case* 3

γ \ $+t$	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·8	1·0	1·2	1·4	1·6	1·8	2·0	3·0
30°	0·66	0·55	0·47	0·40	0·35	0·30	0·27	0·21	0·17	0·14	0·12	0·10	0·09	0·08	0·04
45°	0·98	0·83	0·71	0·61	0·53	0·46	0·41	0·33	0·26	0·22	0·18	0·16	0·13	0·12	0·07
60°	1·27	1·10	0·95	0·83	0·73	0·64	0·57	0·45	0·37	0·30	0·25	0·22	0·19	0·16	0·09
90°	1·80	1·63	1·47	1·33	1·19	1·07	0·97	0·78	0·64	0·52	0·40	0·37	0·31	0·27	0·14
120°	2·21	2·10	2·00	1·91	1·81	1·70	1·58	1·37	1·10	0·90	0·74	0·61	0·50	0·42	0·21
135°	2·35	2·30	2·25	2·22	2·18	2·13	2·06	1·83	1·54	1·24	0·99	0·79	0·65	0·53	0·24
150°	2·46	2·44	2·46	2·51	2·57	2·65	2·71	2·71	2·38	1·86	1·40	1·07	0·83	0·66	0·28

γ \ $-t$	0·2	0·4	0·6	0·7	0·8	0·9	1·0	1·2	1·4	1·6	1·8	2·0	3·0	4·0
30°	0·99	1·61	2·91	4·17	6·53	13·2	∞	5·48	2·25	1·24	0·78	0·53	0·15	0·07
45°	1·40	2·11	3·41	4·59	6·83	13·3	∞	5·65	2·46	1·43	0·92	0·66	0·20	0·10
60°	1·74	2·43	3·65	4·78	6·95	13·4	∞	5·72	2·55	1·52	1·02	0·74	0·24	0·12
90°	2·21	2·79	3·86	4·92	7·03	13·4	∞	5·77	2·62	1·59	1·09	0·81	0·29	0·15
120°	2·48	2·94	3·94	4·97	7·06	13·7	∞	5·78	2·64	1·62	1·12	0·83	0·31	0·16
135°	2·56	2·99	3·96	4·98	7·10	13·4	∞	5·78	2·65	1·63	1·13	0·84	0·31	0·17
150°	2·67	3·01	3·97	4·99	7·07	13·4	∞	5·79	2·71	1·63	1·13	0·85	0·32	0·17

Table 7. *Values of* $(8\pi R/U)E$ *for points on* O_1X_1; $t = x_1/R$. *Case* 3

γ \ $\pm t$	0·0	0·2	0·4	0·6	0·8	1·0	1·2	1·4	1·6	1·8	2·0	3·0	4·0	5·0
30°	0·66	0·64	0·59	0·51	0·45	0·35	0·29	0·24	0·20	0·16	0·14	0·07	0·04	0·03
45°	0·98	0·96	0·90	0·81	0·70	0·58	0·47	0·38	0·31	0·26	0·22	0·10	0·06	0·04
60°	1·27	1·26	1·23	1·17	1·06	0·90	0·72	0·57	0·45	0·37	0·30	0·14	0·08	0·05
90°	1·80	1·80	1·82	1·93	2·34	∞	1·74	1·07	0·76	0·58	0·47	0·20	0·11	0·07
120°	2·21	2·18	2·12	2·02	1·84	1·56	1·25	0·98	0·78	0·63	0·52	0·24	0·14	0·09
135°	2·35	2·31	2·17	1·96	1·69	1·40	1·14	0·94	0·75	0·62	0·52	0·25	0·14	0·09
150°	2·46	2·39	2·19	1·91	1·61	1·32	1·08	0·89	0·73	0·61	0·51	0·25	0·15	0·10

Table 8. *Values of* $(8\pi R/U)\sigma$. *Case* 3

γ \ ϕ/γ	0·0	0·02	0·05	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	0·95	0·98	1·00
30°	∞	12·2	62·0	24·4	12·4	8·52	6·65	5·65	5·11	4·92	5·14	6·31	8·37	12·7	∞
45°	∞	81·6	32·7	16·3	8·29	5·69	4·46	3·79	3·44	3·33	3·49	4·31	5·73	8·75	∞
60°	∞	60·8	24·4	12·2	6·23	4·29	3·37	2·88	2·62	2·55	2·69	3·35	4·47	6·83	∞
90°	∞	40·6	16·2	8·17	4·17	2·89	2·29	1·98	1·83	1·81	1·95	2·48	3·35	5·16	∞
120°	∞	30·4	12·2	6·14	3·15	2·21	1·77	1·56	1·47	1·50	1·67	2·21	3·04	4·76	∞
135°	∞	27·1	10·9	5·46	2·82	1·98	1·61	1·43	1·38	1·43	1·63	2·22	3·12	4·92	∞
150°	∞	24·3	9·75	4·92	2·55	1·81	1·49	1·35	1·32	1·41	1·66	2·36	3·41	5·49	∞
180°	∞	20·3	8·14	4·12	2·17	1·57	1·34	1·27	1·34	1·57	2·17	4·12	8·14	20·3	∞

Points on O_1X_1 $\quad t = x_1/R$.

$$E = (U/\pi^2 R)\sin(\gamma/2)/[(t^2+1)^2(t^4+2t^2\cos 2\gamma+1)]^{\frac{1}{4}} \tag{23}$$

At the centre,

$$E = (U/\pi^2 R)\sin(\gamma/2) \tag{24}$$

Surface density of charge σ.

$$\sigma = (U/4\pi^2 R)\sqrt{(2)}\sin(\gamma/2)/\sin(\phi/2)\sqrt{(\cos\phi-\cos\gamma)} \tag{25}$$

σ possesses the minimum value $(U/2\pi^2 R)\sqrt{(2)}/\sin(\gamma/2)$ when $\cos\phi = (\cos\gamma+1)/2$.

Tables 6, 7 and 8 give the values of $(8\pi R/U)E$ and $(8\pi R/U)\sigma$ for selected values of γ.

THE CAPACITY OF A SPLIT CYLINDRICAL CONDENSER

The capacity of the split cylindrical condenser under consideration is the capacity of the original system, namely, $K'/4\pi K$.

The ratio K'/K may be obtained from tables. Where necessary the following relations facilitate accurate and speedy calculation:

$$\pi K'/K = \ln(1/q) \tag{26}$$

where $q = \epsilon + 2\epsilon^5 + 0(\epsilon^9)$ and ϵ is defined by the relation

$$2\epsilon = (1-\sqrt{k'})/(1+\sqrt{k'}) \tag{27}$$

From these it is easily shown that when $k < 0{\cdot}2$

$$K'/K = (2/\pi)\ln(4/k) \tag{28}$$

On interchanging k and k', we obtain

$$K'/K = \pi/[2\ln(4/k')] \tag{29}$$

Proceeding to the particular cases:

Symmetrical case; $k = (1-\sin\alpha)(1+\sin\alpha)$.

Landen's transformation gives the useful result that K'/K for this modulus is twice K'/K for modulus $k = \cos\alpha$.

Using this property and equations (28) and (29), we obtain the following results:

$$K'/K = \pi\ln(4\operatorname{cosec}\alpha) \qquad 0 < \alpha \leqslant 12°\,30'$$
$$K'/K = 0{\cdot}748 + (2{\cdot}72)10^{-2}\alpha \qquad 15° \leqslant \alpha \leqslant 35°$$
$$K'/K = (4/\pi)\ln(4\sec\alpha) \qquad 78°\,30' \leqslant \alpha \leqslant 90°$$

These give K'/K correct to three significant figures over the range indicated, α being in degrees.

When B_1D_1 *is a diameter*: k is $\tan(\pi/4 - \alpha/2)$

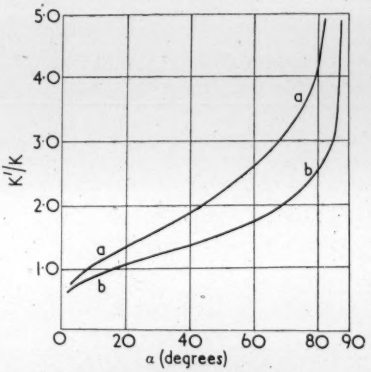


Fig. 4. Graphs of K'/K against α
(*a*) symmetrical case; (*b*) diameter case.

The following relations give K'/K correct to three significant figures over the range indicated.

$$K'/K = 0{\cdot}6963 + (1{\cdot}58)10^{-2}\alpha \qquad 20° \leqslant \alpha \leqslant 45°$$
$$K'/K = (2/\pi)\ln[4\tan(\pi/4+\alpha/2)] \qquad 67° \leqslant \alpha \leqslant 90°$$

Table 9 gives the values of K'/K for a given α in the two

Table 9. *Value of* K'/K *for given* α

α (*degrees*)	2	4	6	8	10	15	20	25	30	35	40	45
Diameter case	0·578	0·663	0·725	0·777	0·822	0·921	1·009	1·091	1·170	1·250	1·330	1·414
Symmetrical case	0·664	0·776	0·863	0·937	1·004	1·155	1·293	1·429	1·559	1·703	1·845	2·000
α (*degrees*)	50	55	60	65	70	75	80	82	84	86	88	90
Diameter case	1·504	1·599	1·709	1·832	1·986	2·168	2·434	2·576	2·738	3·018	3·460	∞
Symmetrical case	2·168	2·351	2·558	2·798	3·092	3·465	3·995	4·276	4·640	5·157	6·039	∞

cases and the graphs of K'/K against α are shown in Fig. 4. The aforementioned linearity is clearly evident.

REFERENCES

(1) THOMSON, J. J. *Recent Researches in Electricity and Magnetism.* § 245, p. 236 (Oxford: Clarendon Press, 1893).

(2) ADAMS, E. P. *Proc. Amer. Phil. Soc.*, p. 251 (1936).

(3) WHITTAKER, and WATSON, G. N. *A Course of Modern Analysis*, § 21.8, p. 485; § 22.42, p. 507 (London: Cambridge University Press, 1946).

Tables giving K'/K tabulated against k^2:

OBERHETTINGER, F., and MAGNUS, W. *Anwendung der Elliptischen Funktionen in Physik und Technik* (Berlin: Springer, 1949).

ATTWOOD, C. *Advanced Five Figure Mathematical Tables* (London: Macmillan and Co. Ltd., 1951).

X-ray determination of crystallinity in deformed natural rubber

By S. C. NYBURG, B.Sc., Ph.D.,* British Rubber Producers' Research Association, 48, Tewin Road, Welwyn Garden City, Herts.

[*Paper first received* 18 *December*, 1953, *and in final form* 22 *April*, 1954]

A vulcanized natural rubber has been deformed in simple extension and in pure shear and the resulting diminution in X-ray intensity of the "amorphous" halo used as a measure of crystallinity. The effect of mode of deformation and attempts to reach "equilibrium" states by heat treatment at constant deformation are described. The values obtained for simple extension are in accord with those reported by Goppel and not with those of Field. The degree of crystallinity is greater in pure shear than in simple extension at the same principal extension ratio. Corrected values for the orientation of crystalline regions have been obtained and a comparison is made between volume changes and X-ray measurements for simple extension.

1. INTRODUCTION

Two previous X-ray determinations of crystallinity in simply-extended natural rubber, those of Field[1] and of Goppel,[2] showed considerable differences in magnitude. A re-determination with an extension to other types of strain has been carried out using a rubber vulcanized with di-*tert.*-butyl peroxide. This is a more suitable material than the conventional sulphur vulcanizates measured by the earlier investigators for two reasons. First, peroxide vulcanization, which is thought to entail molecular chain cross-linking by polyisoprenic carbon atoms themselves, obviates the introduction of foreign material into the rubber and second, the volume change caused by crystallization, which is obscured in sulphur vulcanizates by vacuole formation but appears genuine in peroxide rubbers, has already been measured in these laboratories by Stern.[3]

The theory underlying the determinations is that since the total X-ray intensity from a given specimen is independent of the arrangement of the constituent atoms[4] there will be a compensating diminution in integrated amorphous intensity when an initially amorphous material is caused to show crystalline-type scatter by deformation or in some other way. The degree of crystallinity is then defined as the fractional decrease in integrated amorphous intensity, but this ignores any crystalline regions so small or so disordered that they would give rise to scatter inseparable from the amorphous scatter. Accordingly, it does not necessarily follow that crystallinity determined by X-rays will be the same as that given by other methods which will be differently sensitive to that crystallinity which escapes X-ray measurement.

The amorphous intensity *distribution* round the X-ray beam has been found, by other investigators and in the present work, to be independent of deformation. Accordingly, it is sufficient to measure, say, the diminution in *peak* intensity of the most intense halo at $(\sin\theta)/\lambda = 0{\cdot}10$.

If I_u is the amorphous halo peak intensity of an undeformed, completely amorphous specimen and, other things being equal, I_d the corresponding intensity in the deformed state, the percentage crystallinity is $100(I_u - I_d)/I_u$.

2. MATERIAL AND APPARATUS

Crêpe rubber was lightly milled with the vulcanizing reagent and pressed into smooth sheets 2·5 mm thick. It was extracted with hot acetone for two days after vulcanization, dried in a vacuum and stored in the dark in a vacuum. It had a molecular weight between cross-links, estimated by equilibrium swelling in benzene,[5] of 7100 and an analysis corresponding to 96% rubber hydrocarbon.

The camera, shown in Fig. 1(*a*), was a modified version of the single-crystal gonimeter by Unicam Instruments (Cambridge) Ltd. The collimator was mounted more rigidly than in the standard design and was not moved between exposures. Inside the 3 cm radius film holder, and coaxial with it, was a perspex tube 5 mm thick, to the bottom of which was fitted a slotted steel ring which could be firmly located on the goniometer base. An equatorial zone of the tube was cut away apart from four uprights one of which, central in the beam, held a powder pellet for exposure standardization as in Goppel's measurements. Although the investigation was confined to exposures made at room temperature the inside and outside of the tube were covered with 0·25 mm celluloid sheet to provide insulation should exposures at other temperatures be required. The powder pellet was held in a brass bush and fitted with a thin steel conical screen as shown in Fig. 1(*b*). The whole assembly was rigidly held to the body of the goniometer by an upper circular brass plate bolted to uprights *A*. A semi-cylindrical section, 8·3 cm high, was cut from the original film holder and a new, flanged holder made to fit [Fig. 1(*c*)].

The radiation used was nickel-filtered copper at 45 kV from a CA–6 type sealed-off self-rectified tube. Half the film was screened with 0·31 mm aluminium sheet *B*; the

* Now at Department of Chemistry, University College of North Staffordshire.

difference in intensity on the two halves of the film simulates monchromatic values. Owing to increasing white radiation with target use, the somewhat imperfect monochromatization caused a small drift in calibration during sixty hours' total exposure. Allowance for incoherent scatter was made from the data of Simard and Warren.[6]

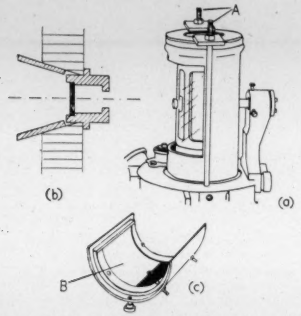


Fig. 1. (*a*) X-ray camera; (*b*) Powder assembly; (*c*) Film holder

The powder line used for standardization was the 4·3 Å reflexion from a pellet of calcium sulphate dihydrate 0·3 mm thick diluted with a four-fold volume of graphite so that it could easily be made of uniform thickness.

Once deformed as described below, the appropriate region of the rubber was compressed between steel rectangles having the channelled-out cross-section shown in Fig. 2. Over these rectangles were placed slotted steel strips *C*

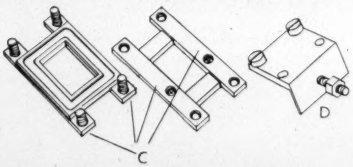


Fig. 2. Specimen assembly

having threaded holes at the ends. Once in place, pressure was applied by a G-clamp which left the holes at the ends of the strips clear. The excess rubber was trimmed away, the two rectangular clamps tightly screwed together and the G-clamp removed. The assembly was then screwed to the support *D* which in turn was screwed firmly into the central stem of the goniometer.

The films (Industrial B by Ilford Ltd.), were photometered with a non-recording photometer (by Hilger and Watts Ltd.). The two types of deformation examined had principal extension ratios α_1 vertical and the plane of the sheet normal to the beam. On the resulting photographs the amorphous halo can be photometered along a radial line inclined at 15° to the equator.

3. PRELIMINARY PROCEDURE

Photographs were corrected for air scatter by taking a number of exposures without rubber specimens and photometering at $(\sin\theta)/\lambda = 0{\cdot}1$. The correction was assumed to hold when the rubber specimens were in place.

The ratio of intensity from the rubber to that from the powder should be independent of intensity and a typical set of results for a specimen about 1 mm thick is shown in Fig. 3. Three such ratios were determined for each specimen.

Both Field and Goppel assumed that the ratio of integrated amorphous intensity to the intensity of the standard powder

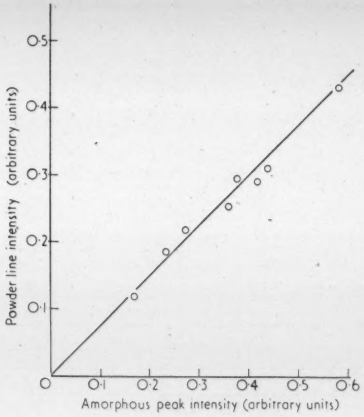


Fig. 3. Initial calibration

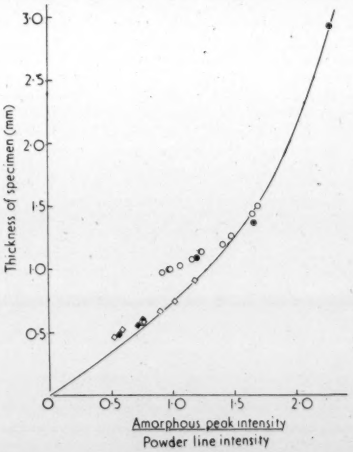


Fig. 4. Intensity ratios as a function of specimen thickness

⊕ = undeformed; ○ = simple extension, unheated; ● = simple extension, heat treated; ◇ = pure shear, unheated; ◆ = pure shear, heat treated.

line would be proportional to the thickness of the specimen. It is easy to show that for a sheet normal to the incident beam this would not be expected to be the case except for very small thicknesses. Moreover, it is safer not to *assume* a form for the ratio but to measure it directly.

Very thin sheets of undeformed rubber, the thickness of which is difficult to measure accurately, were not used for this calibration since deformed specimens, which are sufficiently hard and uniform to give reliable thickness measurements with a gauge, give intensity ratios which converge on the calibration line as deformation is decreased. This, together with the non-linearity of the calibration, is seen in Fig. 4.

4. CRYSTALLINITY INDUCED BY SIMPLE EXTENSION

Simple extension is attained by stretching narrow strips to a principal extension ratio $\alpha_1 = \alpha$; the other extension ratios are then $\alpha_2 = \alpha_3 \simeq \alpha^{-\frac{1}{2}}$. Volume changes are only about 2% at $\alpha_1 = 5{\cdot}0$ so that the thickness, which has in any case to be known, can be used for estimating α_1.

The strips, of measured initial thickness, were stretched by screw-operated clamps to the desired value of α_1 and then fixed vertically at the centre of the rectangles described above. The stretching was normally completed in five minutes. The results, shown in Fig. 5, curve (*a*) are taken from the values in Fig. 4.

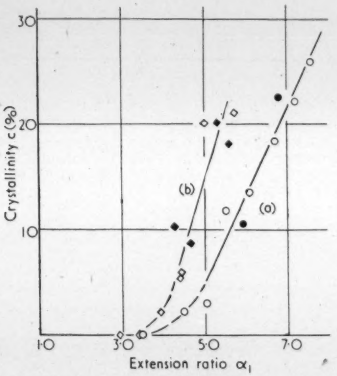


Fig. 5. Dependence of crystallinity on extension for curve *a*, simple extension; curve *b*, pure shear. Key is same as for Fig. 4

It is known that the degree of crystallinity of rubber depends not only upon the strain but upon the route by which the strain is attained. If highly extended specimens are partially relaxed they show hysteresis effects which can be removed by heating. The result of heat treatment is therefore a guide to the closeness of the specimen to its equilibrium crystallinity. Two specimens in simple extension were heated at 90° C (see Section 5) in a vacuum for one hour and allowed to cool for twenty-four hours at room temperature. The results for these are also shown on Fig. 4, the small changes in crystallinity caused by this treatment being in accord with Goppel and Arlman's observations.[7]

5. CRYSTALLINITY INDUCED BY PURE SHEAR

Pure shear is attained when the extension ratios are $\alpha_1 = \alpha$, $\alpha_2 = 1{\cdot}0$ and $\alpha_3 \simeq \alpha^{-1}$. The crystalline type scatter is different from that in simple extension because of "higher orientation."[8]

The undeformed specimens were sheets 65 × 15 mm, the principal stress being applied normal to the longer edge. In addition to the screw-operated clamps used for simple extension, screw-operated clamps mounted at right angles were used to obtain a sideways extension ratio α_2 of unity.

Two deformation routes were studied: (i) that with sideways extension α_2 kept constant at unity throughout the entire deformation ("constant α_2" method) and (ii) that with no sideways stress applied until the principal extension α_1 had reached the required value ("α_2 to unity last"). For values of α_1 above about 5·0, method (ii) brought about irregular yielding of the rubber. Deformation (ii) is arbitrary since the dependence of α_2 on α_1 throughout the deformation will depend upon the initial dimensions of the sheet. The initial dimensions were, however, the same for all specimens.

The specimens deformed by the constant α_2 method have crystallinities which are higher than those in simple extension at the same α_1. [Fig. 5, curve (*b*); data from Fig. 4.] Moreover, these specimens show little alteration in crystallinity on heating at 90° C in a vacuum for one hour and cooling at room temperature for twenty-four hours. 90° C is the highest temperature to which rubber at $\alpha_1 = 5{\cdot}0$ can be heated without rupture occurring.

A few determinations were carried out to study the effect of heat on specimens deformed in pure shear with α_2 to unity last. These determinations were not as extensive or as accurate as those on unheated specimens and only general remarks can be made. After heating a specimen deformed to $\alpha_1 = 5{\cdot}5$ for about thirty minutes and cooling to room temperature there was no immediate attainment of a significantly greater crystallinity. This only appeared on standing for one or two days. On subsequent heating, however, for a further thirty minutes, the attainment of high crystallinity on standing is much more rapid. Successive heat treatments, although promoting rapid attainment of equilibrium on cooling, tend to lower the optimum crystallinity, presumably because of thermal or oxidative degradation of the material. The optimum crystallinity attained by heat treatment was usually only about one half of that attained in pure shear deformation by the α_2-constant method.

6. CRYSTALLITE ORIENTATION

The most usual estimates of tangential angular spread of crystalline-type reflexions are $\phi_{\frac{1}{2}}$ values, that is angles measured from the reflexion centre ($\phi = 0$) to a point where the intensity falls by half, or $\overline{\phi}$ values given by

$$\overline{\phi} = \int_{-180°}^{+180°} I_\phi d\phi / I_0$$

where I_ϕ is intensity at ϕ.

Few quantitative estimates have been made previously. Goppel[7,9] quotes $\phi_{\frac{1}{2}}$ values for a number of rubbers in simple extension without correction for the finite width of the incident beam. A correction was used in the present case based on that due to Jones[10] for θ-broadening, except that instead of a powder, a crystal plate of mica giving a single reflexion at about the same Bragg angle of 14° was used to estimate the corrected $\overline{\phi}_\beta$ of the (200) reflexion (unit cell due

to Bunn[11]). Profiles of both rubber and mica reflexions recorded on an 11·46 cm radius camera were almost Gaussian so that the Gaussian correction curve of Jones was used.

Ten $\bar{\phi}_\beta$ values were determined on heated and unheated specimens in simple extension and in pure shear (constant α_2 deformation) in the range α_1 4·00 to 6·80. All the results were similar in magnitude, lying between 6·7° and 9·1°. The scatter of the points prevented the deduction of definite conclusions regarding the dependence of $\bar{\phi}_\beta$ on α_1. In general, heat treatment appeared to improve $\bar{\phi}_\beta$ in the more highly-extended specimens by about 1·5° in accord with the one observation of Goppel and Arlman[7] for simple extension.

Specimens deformed in pure shear with α_2 to unity last show marked orientation effects which parallel the changes in crystallinity, i.e. on heating, orientation improves but does not reach that attained for constant α_2 deformation.

An objection which might be raised concerning $\bar{\phi}_\beta$ measurements for pure shear is that although the extended (α_1, α_2) plane was mounted normal to the X-ray beam the $\bar{\phi}_\beta$ measurements do not refer to this plane owing to the Bragg condition $2\theta = 14°$. A specimen was therefore mounted with its extended plane set at 83° to the direct beam on the reflexion side. No significant difference in $\bar{\phi}_\beta$ could be detected.*

7. COMPARATIVE VOLUME CHANGES

Volume changes for simply-extended rubbers of this type have been determined by Stern[3] using the method described by Gee, Stern and Treloar.[12] The unit cell of Bunn[11] gives a density for fully crystalline rubber as 1·008 g/cm^3. Taking Stern's value for the amorphous rubber as 0·9184g/cm^3 the relation between percentage crystallinity c and percentage volume change $100\,\Delta V/V$, is

$$c = 11{\cdot}2\,(100\Delta V/V)$$

The crystallinities calculated from volume changes are plotted together with X-ray crystallinity values (from Fig. 4) in Fig. 6. There is a considerable difference between the two sets of results.

The only other reported correlation using Bunn's data is that due to Arlman[16] on undeformed, unvulcanized rubbers which had crystallized on standing. The crystallinities calculated from density measurements were in exact accord with those determined by the X-ray method.

8. DISCUSSION

The dependence of crystallinity and of orientation on α_1 in simple extension is in full agreement with Goppel's[2] determinations for comparable rubbers. Crystallinity values of Field[1] must be considered too high, probably as a result of unsatisfactory exposure standardization.

No explanation is offered for the greater crystallinity in pure shear than in simple extension at the same α_1. The comparatively small difference in orientation for the two deformations is unexpected. A rapidly increasing dependence of $\bar{\phi}_\beta$ on α_2 is to be anticipated for α_2 values greater than unity since in two-dimensional extension ($\alpha_1 = \alpha_2 = \alpha$; $\alpha_3 \simeq \alpha^{-2}$), $\bar{\phi}_\beta$ is 360°.

The effect of heat treatment on simply-extended rubber and on specimens deformed in pure shear by the constant α_2 method could be due to their being in a state of genuine equilibrium crystallinity, and the behaviour of heated specimens deformed in pure shear by the alternative route tends to support this view.

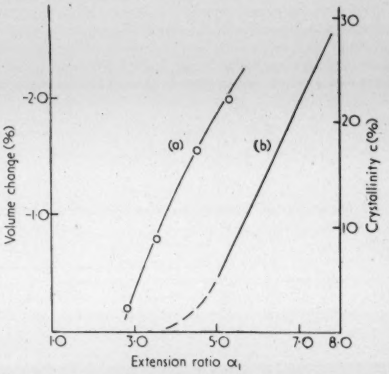


Fig. 6. Comparative volume and X-ray measurements (*a*) Stern's results; (*b*) X-ray results

The irregular yielding of specimens subject to increasing α_2 could be simply explained in terms of disruption of the initial crystallinity.

It seems unlikely that the large difference in crystallinity as estimated from volume changes could be attributed to error in the assumed density for fully crystalline rubber. The difference may be due to the presence of small or disordered regions of crystallization (see Section 1), but more correlative determinations are required to settle this point.

ACKNOWLEDGEMENTS

I am grateful to Dr. J. Stern for use of his unpublished data.

This work forms part of a programme of research undertaken by the Board of the British Rubber Producers' Research Association.

REFERENCES

(1) Field, J. E. *J. Appl. Phys.*, **12**, p. 23 (1941).
(2) Goppel, J. M. *Appl. Sci. Res.* (*A*), **1**, p. 3 (1949).
(3) Stern, J. Unpublished results (1951).
(4) Compton, A. H., and Allison, S. K. *X-Rays in Theory and Experiment*, p. 189 (New York: Van Nostrand and Co. Inc., 1935).
(5) Gee, G. *Trans Faraday Soc.*, **42**, p. 585 (1946).
(6) Simard, G. L., and Warren, B. E. *J. Amer. Chem. Soc.*, **58**, p. 507 (1936).
(7) Goppel, J. M., and Arlman, J. J. *Appl. Sci. Res.* (*A*), **1**, p. 462 (1949).
(8) Mark, H., and v. Susich, G. *Kolloid. Z.*, **46**, p. 11 (1928).
(9) Goppel, J. M. *Appl. Sci. Res.* (*A*), **1**, p. 18 (1949).
(10) Jones, F. W. *Proc. Roy. Soc. A*, **166**, p. 16 (1938).
(11) Bunn, C. W. *Proc. Roy. Soc. A*, **180**, p. 40 (1942).
(12) Gee, G., Stern, J., and Treloar, L. R. G. *Trans Faraday Soc.*, **46**, p. 1101 (1950).
(13) Arlman, J. J. *Appl. Sci. Res.* (*A*), **1**, p. 347 (1949).

* *Note added in proof:* This lack of detectable difference is now known to be due to the relatively poor extent of higher orientation in pure shear. [Nyburg, *Acta Cryst.*, **7**, p. 385 (1954).]

An improved technique for the micro-electrophoresis of oil drops

By D. A. Smith, M.Sc., A.R.I.C., A.I.R.I.,* National College of Rubber Technology, London, N.7

[*Paper received* 7 *January*, 1954]

A new technique is described for preparing hydrocarbon dispersions by the "steam-jet" method without risk of atmospheric contamination. A modified cylindrical micro-electrophoresis cell in conjunction with a conventional apparatus is used to establish conditions for reproducible mobility measurements. Measurements were made between 2½ and 7 h after the preparation of the dispersion using, as dispersion medium, sodium chloride or pyroborate solution at a controlled pH. Values obtained for the mobilities of cyclohexane and *n*-hexane at 24·9° C in 0·01N Na^+ were, at pH 9·0, 2·94 and 3·52 μ per sec per V per cm respectively; at pH 11, 3·22 and 3·90 μ per sec per V per cm; and in 0·1N Na^+ at pH 9·0, cyclohexane gave the value 2·69 μ per sec per V per cm.

Application of the technique of micro-electrophoresis to the investigation of the surface properties of dispersed hydrocarbon droplets has been discussed by a number of investigators.[1–4] Two types of micro-electrophoresis cell have proved popular, the rectangular cell of Abramson[5] and the cylindrical cell of Mattson.[6] Modifications of these have been described by Smith and Lisse,[7] Bradbury and Jordan,[8] van Gils and Kruyt[9] and, latterly, by Douglas[10] and by Alexander and Saggers.[11] A semi-micro-electrophoresis apparatus, which is particularly valuable for the investigation of dispersions in very dilute electrolyte solutions, has been developed by Jordan and Taylor.[12] Three methods have been used to prepare dispersions of oils in aqueous media[13]; use of the colloid mill introduces a grave risk of dispersion contamination and there is evidence that the injection of a solution of an oil in ethyl alcohol into boiling water allows sufficient alcohol to remain adsorbed on the oil droplets to alter radically the electrophoretic properties of the dispersion.[12] In the present research, the multi-steam-jet method described by Douglas has been modified to allow the manipulation of dispersions with the minimum risk of contamination by surface-active impurities or by atmospheric carbon dioxide.

DESCRIPTION OF APPARATUS

The apparatus was based on that of Alexander and Saggers[11] and consisted of a trough through which was circulated water from a thermostat bath and in which the electrophoresis cell was supported. The cell was illuminated from above and observed through the side of the trough by means of a travelling microscope.

The electrophoresis cell (Fig. 1) was modified from Mattson's original design[6] which takes advantage of the simple equations for hydrodynamic flow within a cylindrical tube. Replacement of the single filling and drain tubes by dual tubes connecting the central filler tube to either the electrode compartment or the body of the cell by way of a hollow-bore Y-stopcock facilitated the various operations of cleaning and filling the cell.

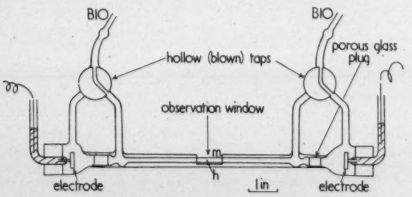


Fig. 1. Micro-electrophoresis cell

The electrodes were of the silver/silver chloride (reversible) variety and were fixed in the electrode compartments by means of rubber bungs which had been conditioned by immersion in distilled water for several months. The use of ungreased ground-glass joints in place of bungs was found to be unsatisfactory, since sticking of the joints followed long immersion in water or salt solutions. The electrode compartments were separated from the central portion of the cell by means of sintered glass plates fused into the tube; a No. 2 sinter plate was found satisfactory. This enabled different solutions to be used in the cell and in the electrode compartments and made it possible to change the dispersion under observation by flushing through the cell from the reservoir. Thus it was possible to detect accidental contamination caused by diffusion of surface-active impurities from the electrode compartments during the course of a run. No such contamination was observed during an extended series of experiments.

The dispersing unit (Fig. 2) was constructed of Pyrex glass and consisted of a reservoir of 250 ml capacity fitted with a thermometer pocket and drain tap. The lid of the reservoir was in the form of a specially ground cone joint through which passed a filler tube and gas exit tube fitted with a hollow-bore tap. A wider tube passed through the centre of the lid beneath which was supported a circular glass tube carrying five jets of 0·2 mm average diameter.

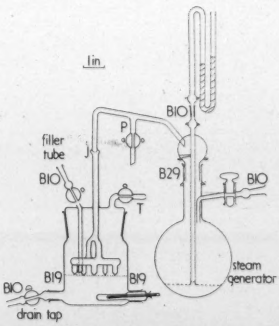


Fig. 2. Dispersing unit

* Now at Atomic Weapons Research Establishment, Aldermaston, Berkshire.

The upper end of the centre tube was connected through a spherical joint to a glass steam generator, a 500 ml Pyrex flask with fluted sides and fitted with a mercury manometer, an inlet tube for charging the boiler and a pressure release tap. The fluted sides made possible a very rapid conversion of water into steam, without bumping, when heated by a gas burner. All glass taps were of the hollow type. They were not greased, neither was grease applied to any joints. (An exception was made in the case of the pressure-release tap which was lightly greased for reasons of safety; it was not in the direct line from the boiler to the dispersing chamber.)

All glass apparatus (the dispersing chamber, steam generator, pH and conductivity cells and the electrophoresis cell) was cleaned as follows. The apparatus was soaked in a warm chromic-sulphuric acid mixture for two hours and washed thoroughly in tap water. It was rinsed twice in distilled water and allowed to soak in distilled water overnight. It was then emptied and rinsed three times with conductivity water prepared by several distillations of "ion-exchange" water in a counter-current of pure nitrogen.

(The chromic acid treatment of glassware for electrophoresis measurements has been criticized by Davies and Holliday.[14] They quote evidence that glass, after exposure to neutral or acid solutions, loses alkali at the surface, becoming covered with a thin layer of gelatinous silicic acid which can in turn be attacked by strong alkali or silver salts. In order to test if this were the case in these experiments, the electrophoresis cell was rinsed several times with N/5 sodium hydroxide, after exposure to chromic acid, and then rinsed with water in the usual way. No significant change in mobility values was observed. The alternative washing technique of Davies and Holliday, i.e. sodium thiosulphate followed by dilute hydrochloric acid, appears less likely than is chromic acid to destroy traces of surface-active materials which may affect the mobility values.)

PREPARATION OF A SAMPLE FOR ELECTROPHORESIS

The steam generator was flushed with purified nitrogen; 50 ml of conductivity water ($\kappa = 1 \times 10^{-6}$ mho/cm) was introduced through the filler tap from an automatic burette. Sufficient of the purified solid salt (0·0877 g NaCl or 0·2861 g $Na_2B_4O_7.3H_2O$) was introduced into the dispersing chamber to make an exactly N/100 solution. The chamber was flushed with nitrogen and exactly 150 ml conductivity water was added from the grade A automatic burette through the filler tube. The salt dissolved readily. In the sodium chloride experiments, the pH was adjusted to the desired value by the addition of a few drops of N/20 sodium hydroxide solution prepared under nitrogen in a polyethylene wash-bottle, the nozzle of which was fitted with a B10 ground-glass cone for ready attachment to the dispersion chamber filler tube. Samples were withdrawn from the drain tap into special micro-pH and conductivity cells. These were allowed to stand in the thermostat for ten minutes before the equilibrium reading was taken. No pH measurements were necessary in N/100 sodium pyroborate without added sodium hydroxide since the pH value of this solution had been determined accurately in an earlier experiment as 9·00 ± 0·02 pH units at 25° C, using as standard a N/20 borate buffer (pH = 9·14 at 25° C).

A few drops of the hydrocarbon to be dispersed were added down the filler tube and floated on the surface of the salt solution which was run out through the drain tap until the jets were in the surface of the liquid. Steam was generated in the flask and was at first allowed to go to waste through the pressure release tap *P* (Fig. 2). The gas exit tap *T* was then opened and *P* closed. Dispersion of the oil was effected in about fifteen seconds; then *P* was opened, *T* closed and the dispersing chamber disconnected from the steam generator at the joint *J* which was stoppered to prevent entry of air. The pressure of steam during dispersion was some 6–8 cm of mercury and the volume condensing in the liquid negligible (less than 1 ml). After the dispersion had been prepared, the electrophoresis cell was connected to the drain tap of the dispersion chamber, rinsed twice with the dispersion, and then rinsed continuously until the volume of dispersion which had passed through was equal to about five times that of the cell. The cell and drain taps were closed and the cell disconnected; the rubber corks of the electrode compartments were given a coating of poly-methyl methacrylate cement to prevent leakage of current during the measurements and allowed to dry.

MEASUREMENT OF ELECTROPHORETIC VELOCITY

After filling, the cell was allowed to stand for thirty minutes in the thermostat trough to attain temperature equilibrium. This temperature was 24·90 ± 0·05° C. A potential of about 200 V was applied across the cell and the current adjusted to a suitable value by means of a variable resistance. The microscope was focused on the "stationary level" by the method of Alexander and Johnson.[15] This level had been determined previously by plotting particle velocity against distance from the inner wall of the cell and multiplying the radius, determined thereby, by $(1 - \sqrt{2}/2)$ or 0·294.

The method of Powney and Wood[3] was used in timing the particles. In an experiment, twenty particles were each timed when travelling between marks in the microscope eyepiece scale; with the current reversed, a further twenty particles were each timed in the reverse direction, making a total of forty readings to be averaged for each velocity measurement. In addition, the microscope was refocused on the stationary level after each group of ten readings in order to eliminate any bias due to inaccurate focusing. Care was taken to reverse the current between consecutive readings in order to minimize ionic diffusion from the electrode compartments into the body of the cell. (The mean particle velocities in each direction were identical within experimental limits.) In order to avoid surface-conductivity corrections, it was necessary to make measurements only on hydrocarbon droplets of relatively large diameter.[16,17] However, Mooney has shown that the electrophoretic velocity shows some dependence on drop size for diameters greater than 20 μ.[2] Droplets of about 5 μ diameter were easily selected in the eyepiece scale when the dispersion was new, but, after several hours, it became difficult to make measurements, since such large particles showed a greater tendency to cream to the surface than the smaller ones. As a result, particles had to be selected for timing over a fairly large focus range and timings on old dispersions showed a greater scatter than those on fresh preparations. Measurements on very fresh dispersions gave lower velocity values than those obtained between 2½ and 7 h after dispersion; these low values were not exactly reproducible.

CALCULATION OF ELECTROPHORETIC MOBILITY

The electrophoretic mobility (u) was evaluated as the particle velocity per unit field strength by dividing the mean

particle velocity by the potential gradient. The latter was obtained from the expression:

$$\text{potential gradient} = A/c\kappa \quad \text{V/cm}$$

where A = current passing in amperes
c = cross-section of cell in cm^2
κ = specific conductivity of dispersion in mhos/cm.

Mobility values are expressed in units of μ per sec per V per cm towards the anode. This method of calculating the potential gradient is preferred since contact potentials at the electrodes do not introduce errors.

Since the potential gradient is determinate to within 0·5%, the chief error lies in the measurement of particle velocity. An estimate, based on the ranges of particle velocities from a number of experiments, gives ±3% for a cell of given radius. Absolute values may be subject to a further (constant) error due to the difficulties inherent in the determination of the optical radius of the cell.[15]

THE INVESTIGATION OF FACTORS WHICH INFLUENCE MOBILITY

Four factors have received special attention.

(1) The time elapsing between the preparation of the dispersion and the measurements of particle velocity influences the mean value of the velocity (Fig. 3). Experiments on

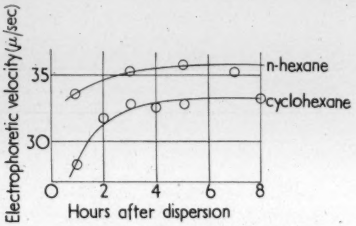


Fig. 3. Relationship between age of dispersion and velocity (typical curves)

cyclohexane and on *n*-hexane showed that the mean particle velocity achieved a constant value about 2–3 h after dispersion and remained constant up to 7–9 h after dispersion, the limiting time for accurate measurements. During the first two hours after dispersion, it was usual to observe a rise in the electrophoretic velocity. This was attributed either to a disturbance of the liquid in the cell by the upward movement of large oil drops (a gravitational effect), or to some delay in the development of the surface charge on the drop, due, perhaps, to a slow orientation of hydroxyl ions and hydrocarbon molecules at the oil/water interface. The reproducibility of mobility values between replicate experiments precluded the possibility of accidental contamination.

(2) Variation of the nature and concentration of the salt dissolved in the aqueous phase caused little variation in the mobility values for a given hydrocarbon. Substitution of sodium pyroborate for sodium chloride caused no detectable variation, suggesting that neither the chloride nor the borate ion is adsorbed in the "inner Helmholtz plane".[18] A ten-fold increase in salt concentration caused a decrease in mobility of about 10% (at fixed pH 9·0) due to the reduction in thickness of the ionic double layer. This confirmed the use of 0·01N Na^+ salts as aqueous media for electrophoretic measurements in order to reduce the thickness of the double layer to a small value which was sensibly independent of small changes in salt concentration.

(3) The influence of pH on mobility (Fig. 4) was found to be similar to that described by Carruthers[13] and was taken as evidence that the hydrocarbon surface becomes saturated with hydroxyl ions at high pH values.

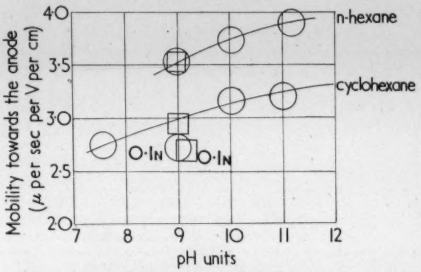


Fig. 4. The influence of salt concentration and of pH on mobility. Concentration 0·01 N except where marked

○ = NaCl medium; □ = $Na_2B_4O_73H_2O$ medium.

(4) The effect of potential gradient on mobility in the range 6–14 V/cm was not greater than the inaccuracies involved in the determination of particle velocity (about ±3%). This is in agreement with the work of Douglas.[10]

It was concluded that a reproducible value of the electrophoretic mobility of hydrocarbon droplets could be obtained by measuring their velocity under a known potential gradient, provided that measurements were made between 2½ and 7 h after dispersion, the salt solution was 0·01N in sodium chloride or pyroborate, and the pH of the dispersion was determinate to within 0·1 unit. The temperature was kept constant to within 0·1° C.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. H. W. Douglas of the University of Liverpool, Dr. W. Gerrard of the Northern Polytechnic, London, and Dr. F. H. Cotton of the National College of Rubber Technology, for encouragement and advice, and the Salters' Company and the Institution of the Rubber Industry for scholarships, during the tenure of which the research was completed.

REFERENCES

(1) Ellis, R. *Z. Phys. Chem. A*, **78**, p. 321 (1912).
(2) Mooney, M. *Phys. Rev.*, **23**, p. 396 (1924).
(3) Powney, J., and Wood, L. J. *Trans Faraday Soc.*, **36**, p. 57 (1940).
(4) Douglas, H. W. *Trans Faraday Soc.*, **46**, p. 1082 (1950).
(5) Abramson, H. A. *J. Gen. Physiol.*, **12**, p. 469 (1929).
(6) Mattson, S. *J. Phys. Chem.*, **32**, p. 1532 (1928); **37**, p. 223 (1933).
(7) Smith, M. E., and Lisse, M. W. *J. Phys. Chem.*, **40**, p. 399 (1936).
(8) Bradbury, F. R., and Jordan, D. O. *Biochem. J.*, **36**, p. 287 (1942).
(9) van Gils, G. E., and Kruyt, H. R. *Kolloid-Beihefte*, **45**, p. 60 (1936).

(10) DOUGLAS, H. W. *J. Sci. Instrum.*, **24**, p. 103 (1947).
(11) ALEXANDER, A. E., and SAGGERS, L. *J. Sci. Instrum.*, **25**, p. 374 (1948).
(12) JORDAN, D. O., and TAYLOR, A. J. *Trans Faraday Soc.*, **48**, p. 346 (1952).
(13) CARRUTHERS, J. C. *Trans Faraday Soc.*, **34**, p. 300 (1938).
(14) DAVIES, K. N., and HOLLIDAY, A. K. *Trans Faraday Soc.*, **48**, pp. 1061, 1066 (1952).
(15) ALEXANDER, A. E., and JOHNSON, P. *Colloid Science*, Vol. I, p. 313 (London: Oxford University Press, 1949).
(16) OVERBEEK, J. TH. G. VON. *Kolloid-Beihefte*, **54**, p. 287 (1943).
(17) OVERBEEK, J. TH. G. VON. *Advances in Colloid Science*, Vol. III (New York: Interscience Publishers Inc., 1950).
(18) GRAHAME, D. C. *Chem. Rev.*, **41**, p. 441 (1947).

The elliptic surface wave

By A. E. KARBOWIAK, Ph.D., University College, London*

[*Paper first received* 17 *November*, 1953, *and in final form* 7 *January*, 1954]

In this paper formulae are derived for the surface wave of the elliptic cylinder. The mode described is an *E*-mode that possesses the evanescent structure in the transverse plane of the cylinder. The elliptic surface wave is shown to be similar in some respects to the surface wave associated with single wire transmission lines of circular symmetry to which form the elliptic surface wave degenerates when the eccentricity of the ellipse tends to zero. In a separate section a slightly deformed circular guide is discussed in detail and it is shown that the performance of a circular guide is substantially unaffected by slight deformation of its geometry. The last section of the paper deals with different types of elliptic guides that are likely to be encountered. The elliptic dielectric rod waveguide and a corrugated elliptical guide are typical examples, while the homogeneous metal guide and the dielectric coated perfectly conducting elliptical rod are analysed in detail.

1. INTRODUCTION

Single wire transmission lines have recently received a great deal of attention,[1,2] but the history of the subject dates back to the beginning of this century when Sommerfeld[3] showed theoretically that a single straight metal wire of finite conductivity and of circular cross-section can act as a guide for electromagnetic energy. A few years later Harms[4] showed that a perfectly conducting wire when coated with a layer of dielectric will support a wave similar in character to that discovered by Sommerfeld. About the same time, Hondros and Debye[5] studied waves of similar nature but guided by dielectric rods of circular cross-section. Further, a uniformly corrugated metal rod of suitable proportions can act as a guide for electromagnetic energy,[6] and the wave guided by it is similar to the wave discussed by Sommerfeld.

The Sommerfeld's single wire wave is an example of an *axial* cylindrical surface wave, in particular of circular symmetry. This is a wave[2] that propagates without radiation along the *axis* on the outside of a cylindrical guide of circular cross-section and is characterized by the absence of the axial component of the magnetic field vector (*E*-mode). In the plane normal to the axis of the guide the wave is *evanescent* in nature (i.e. all field quantities decay monotonically with radial distance) and this is a unique feature of this wave mode, which shall be denoted by $E_{\times\bullet}$. In this notation the $\times$ signifies the evanescent nature of the wave in the radial direction and $\bullet$ stresses the fact that the wave is propagated in the axial direction (*Z*-direction).

The performance of a single wire transmission line is governed by its decay coefficient and the larger the decay coefficient, the more the field of the wave is confined to the immediate neighbourhood of the guide. Further, the value of the decay coefficient is entirely a function of the guide radius and its surface impedance but the actual physical structure of the guide is irrelevant. Thus, for example, the surface waves, one on the outside of the dielectric rod, the other on the outside of a metal wire, coated with a thin layer of dielectric are identical, provided that the outside diameters of the two guides and their surface impedances are equal.

Apart from the surface wave of the circular cylinder the plane surface wave associated with the name of Zenneck,[7] was the only other type of surface wave that was studied in greater detail.

In this paper an attempt is made at formulating the basic theory of the elliptic surface wave, and this wave arises for example in the study of deformed circular surface waveguides.

2. WAVE EQUATION IN ELLIPTICAL CO-ORDINATES

In order to treat problems concerning cylinders of elliptic section it is necessary to use the elliptic cylinder co-ordinates

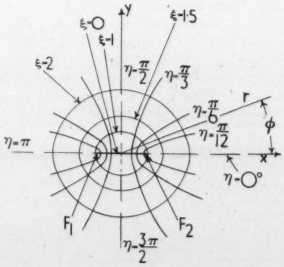


Fig. 1. The elliptical co-ordinates system

ξ, η, z (Fig. 1). With each value of the co-ordinate ξ we associate an ellipse of the family

$$\frac{x^2}{\cosh^2 \xi} + \frac{y^2}{\sinh^2 \xi} = c^2 \qquad (1)$$

* Now at Standard Telecommunication Laboratories Ltd., Enfield, Middlesex.

and with each value of the co-ordinate η we associate a hyperbola of the family

$$\frac{x^2}{\cos^2\eta} - \frac{y^2}{\sin^2\eta} = c^2 \tag{2}$$

Equations (1) and (2) represent a system of orthogonal and confocal curves as shown in Fig. 1. The constant c is equal to half the interfocal length F_1F_2 in Fig. 1. The variable z is a rectilinear co-ordinate and runs parallel to the axis of the elliptic cylinder (Fig. 2).

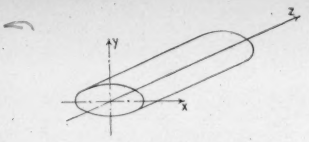


Fig. 2. The elliptical guide

It is assumed that all field quantities are proportional to $\exp(-i\beta z)$ and that the sinusoidal time factor, $\exp(i\omega t)$, is implied. The wave equation for E-modes then becomes:

$$(\nabla_z^2 + k_0^2)E_z = 0 \tag{3}$$

where ∇_z^2 is the two-dimensional Laplacian operator in the plane perpendicular to the axis z.

In the rationalized M.K.S. system of units, which we shall use throughout, we have

$$k^2 = \left(\frac{2\pi}{\lambda_0}\right)^2 = \omega^2\mu_0\kappa_0 = \beta^2 + k_0^2 \tag{4}$$

where μ_0, κ_0 are permeability and permittivity of free-space.

Using the definition of the co-ordinate system (1)–(2), the wave equation (3) becomes:

$$\frac{\partial^2 E_z}{\partial\xi^2} + \frac{\partial^2 E_z}{\partial\eta^2} + \frac{k_0^2c^2}{2}(\cosh 2\xi - \cos 2\eta)E_z \tag{5}$$

Through the substitution

$$E_z(\xi, \eta) = \Xi(\xi).\Lambda(\eta) \tag{6}$$

Equation (6) can be separated into two ordinary differential equations

$$\frac{d^2\Xi}{d\xi^2} - (a - 2q_0.\cosh 2\xi)\Xi = 0 \tag{7}$$

and

$$\frac{d^2\Lambda}{d\eta^2} + (a - 2q_0.\cos 2\eta)\Lambda = 0 \tag{8}$$

In these equations a is a separation constant and is the eigen value of the system (7)–(8) and

$$q_0 = (k_0c/2)^2 \tag{9}$$

The last substitution was made to bring the differential equations to the canonical form as given by equations (7) and (8). The equations (7) and (8) are the Mathieu and associated Mathieu equations respectively. There are several forms of solution to these equations[8] but in the present applications we shall require that the function H shall be periodic in η and then we have as the only solution to equation (8)

$$\Lambda(\eta) = A_1.ce_m(\eta, q_0) \tag{10}$$

or

$$\Lambda(\eta) = A_2.se_m(\eta, q_0) \tag{11}$$

which are the even and odd periodic Mathieu functions of order m. Each of these solutions are associated with a different value of the characteristic number a, and consequently to each solution of equation (8) there will correspond, for the same value of a, a complete solution to equation (7).

In any finite region of ξ the complete solution to equation (7) is

$$\Xi(\xi) = B_1.Ce_m(\xi, q_0) + B_2.Fey_m(\xi, q_0) \tag{12}$$

or

$$\Xi(\xi) = C_1.Se_m(\xi, q_0) + C_2.Gey_m(\xi, q_0) \tag{13}$$

where Ce_m, Se_m, Fey_m, Gey_m are different associated Mathieu functions.[9]

Since the functions Fey_m and Gey_m have each a singularity at the origin, in regions containing the origin the function Ce_m or Se_m is alone the complete solution to equation (7).

The remaining field components are obtained by substituting equation (6) into:

$$\left.\begin{aligned} H_\xi &= \frac{i\omega\kappa_0}{k_1^2}\cdot\frac{1}{h}\cdot\frac{\partial E_z}{\partial\eta} \\ H_\eta &= -\frac{i\omega\kappa_0}{k_1^2}\cdot\frac{1}{h}\cdot\frac{\partial E_z}{\partial\xi} \\ E_\xi &= -\frac{i\beta}{k_1^2}\cdot\frac{1}{h}\cdot\frac{\partial E_z}{\partial\xi} \\ E_\eta &= -\frac{i\beta}{k_1^2}\cdot\frac{1}{h}\cdot\frac{\partial E_z}{\partial\eta} \end{aligned}\right\} \tag{14}$$

where

$$h = c\sqrt{(\cosh^2\xi - \cos^2\eta)}$$

$$= \frac{c}{\sqrt{2}}\sqrt{(\cosh 2\xi - \cos 2\eta)} \tag{15}$$

3. THE NATURE OF THE ELLIPTIC SURFACE WAVE: THE PRINCIPAL E-MODE

Our problem is to investigate in the space outside a single elliptical guide (Fig. 2) the possible existence of an E-mode that possesses the required evanescent structure in the radial direction ξ. The solution corresponding to $m = 0$ [equations (10), (11)] we shall call the principal mode, which is the subject of this investigation.

It will be convenient, here, to make the substitution

$$k_1^2 = -u^2 \tag{16}$$

so that the wave numbers are connected by

$$u^2 = \beta^2 - k^2 \tag{17}$$

Thus

$$\Lambda = A_1ce_0(\eta, -q) \tag{18}$$

where

$$q = (uc/2)^2 \tag{19}$$

For the radial function $\Xi(\xi)$ it will be convenient to have a function akin to the Hankel function $H_0^{(1)}(x)$. This function is defined by

$$Me_0^{(1)}(x, q) = Ce_0(x, q) + iFey_0(x, q) \tag{20}$$

which is analogous to

$$H_0^{(1)}(x) = J_0(x) + iY_0(x)$$

In the present problem the function $Me_0^{(1)}$ alone is a complete solution to equation (7). Thus the field of the elliptic surface wave is derived from

$$E_z = A.Me_0^{(1)}(\xi, -q).ce_0(\eta, -q).\exp(i\beta z) \tag{21}$$

The angular function $ce_0(\eta, -q)$ can be expanded in a Fourier series[9,10]:

$$ce_0(\eta, -q) = \sum_{m=0}^{\infty} (-1)^m . A_{2m} . \cos(2m\eta) \qquad (22)$$

This function is shown plotted in Fig. 3 for purely real values of q using tabulated values of the function.[9,10]

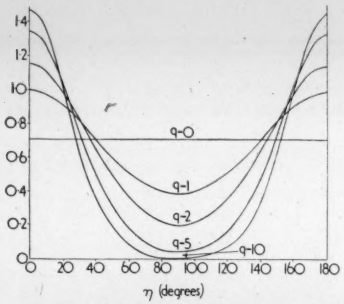


Fig. 3. The Mathieu function $ce_0(\xi, -q)$

Similarly the radial function can be expanded in a series of Hankel functions.[9]

Thus

$$Me_0^{(1)}(\xi, -q) = \frac{ce_0(0, q)}{A_0} \sum_{m=0}^{\infty} A_{2m} . H_{2m}^{(1)}(i2\sqrt{q} \cosh \xi) \qquad (23)$$

Alternatively the function $Me_0^{(1)}$ may be expanded in a Bessel functions product series.[9] Thus:

$$Me_0^{(1)}(\xi, q) = \frac{p_0}{A_0} \sum_{r=0}^{\infty} (-1)^r A_{2r} J_r(v_1) H_r^{(1)}(v_2) \qquad (24)$$

where

$$\left.\begin{aligned} p_0 &= \text{a constant} \\ v_1 &= \sqrt{q} . e^{-\xi} \\ v_2 &= \sqrt{q} . e^{\xi} \end{aligned}\right\} \qquad (25)$$

The function $Me_0^{(1)}$ does not seem to be tabulated in the literature but tables[10,11] of coefficients A_{2m} can be used for its computation and the function so calculated is shown plotted in Fig. 4.

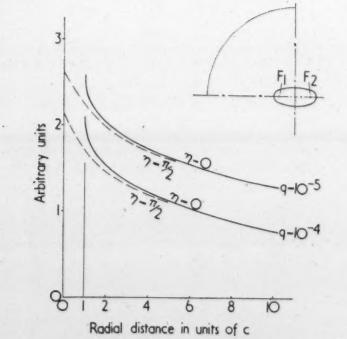


Fig. 4. $Me_0^{(1)}(\xi, -q)$ as a function of r

The co-ordinates ξ and η are related to the circular polar co-ordinates by (Fig. 1):

$$\left.\begin{aligned} r.\cos\phi &= c.\cosh\xi.\cos\eta \\ r.\sin\phi &= c.\sinh\xi.\sin\eta \end{aligned}\right\} \qquad (26)$$

Thus

$$r = c\sqrt{(\cosh^2\xi - \sin^2\eta)} = \frac{c}{\sqrt{2}}\sqrt{(\cosh 2\xi + \cos 2\eta)} \qquad (27)$$

and

$$\tan\phi = \tanh\xi . \tan\eta \qquad (28)$$

Consequently for large values of ξ we have

$$r \to c.\cosh\xi \qquad (29)$$

and the hyperbolic angle η becomes the circular angle ϕ. With this substitution and using asymptotic expansion to Hankel functions in equation (23) it is easy to show that for large values of ξ

$$Me_0^{(1)}(\xi, -q) \simeq \frac{ce_0(0, q) . ce_0\left(\frac{\pi}{2}, q\right)}{A_0} \sqrt{\left(\frac{2}{iur}\right)} . \exp\left(-ur - i\frac{\pi}{4}\right) \qquad (30)$$

and hence

$$E_z \simeq \frac{B}{\sqrt{r}} \exp(-ur) . ce_0(\phi, -q) . \exp(-i\beta z) \qquad (31)$$

where B is a constant.

We note that although the wave will never degenerate into the surface wave of a circular cylinder, however large r might be, yet the wave is evanescent with respect to r and hence by definition, equation (21) represents a surface wave provided u lies in the fourth quadrant of the Argand diagram.

The remaining field components of the surface wave follow from equations (14) and (21). Thus (Figs. 5 and 6):

$$\left.\begin{aligned} H_\xi &= -\frac{i\omega\kappa_0}{u^2} \cdot \frac{A}{h} \cdot Me_0^{(1)}(\xi, -q) \cdot \frac{\partial}{\partial\eta}[ce_0(\eta, -q)] \\ H_\eta &= \frac{i\omega\kappa_0}{u^2} \cdot \frac{A}{h} \cdot \frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi, -q)] \cdot ce_0(\eta, -q) \\ E_\xi &= \frac{i\beta}{u^2} \cdot \frac{A}{h} \cdot \frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi, -q)] \cdot ce_0(\eta, -q) \\ E_\eta &= \frac{i\beta}{u^2} \cdot \frac{A}{h} \cdot Me_0^{(1)}(\xi, -q) \cdot \frac{\partial}{\partial\eta}[ce_0(\eta, -q)] \end{aligned}\right\} \qquad (32)$$

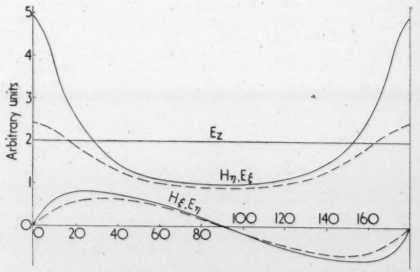


Fig. 5. Field components as a function of η

——— $\xi = 0{\cdot}2$
- - - - $\xi = 0{\cdot}4$ } for $q \ll 1$

And by definition of the guide surface impedance[2,8] we have:

$$Z_s = \frac{E_z}{H_\eta}\bigg|_{\xi=\xi_0} = h_{\xi_0}\frac{u^2}{i\omega\kappa_0}\frac{Me_0^{(1)}(\xi_0,-q)}{\frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi,-q)]_{\xi=\xi_0}} \tag{33}$$

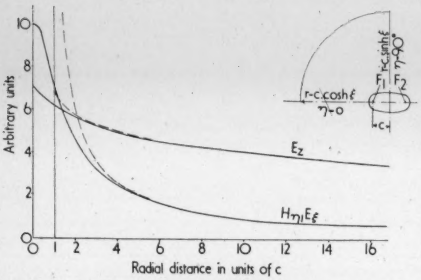


Fig. 6. Field components of the elliptic surface wave as a function of radial distance r

$$q = \left(\frac{uc}{2}\right)^2 = 10^{-6}$$

------ for $\eta = 0°$; ——— for $\eta = 90°$.

We observe that the guide surface impedance of an elliptic guide (Fig. 7)—in contrast to circular guide—is not a constant but is a function of the angle η [equation (15)]. The surface impedance is a maximum for $\eta = 90°$ and 270°.

$$(z_s)_{max} = c.\cosh\xi_0\frac{u^2}{i\omega\kappa_0}\frac{Me_0^{(1)}(\xi_0,-q)}{\frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi,-q)]_{\xi=\xi_0}} \tag{34}$$

and a minimum for $\eta = 0°$ and 180°

$$(z_s)_{min} = c.\sinh\xi_0\frac{u^2}{i\omega\kappa_0}\frac{Me_0^{(1)}(\xi_0,-q)}{\frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi,-q)]_{\xi=\xi_0}} \tag{35}$$

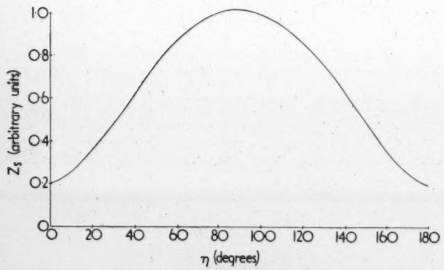


Fig. 7. Z_s as a function of η

Moreover as the guide dimension ξ_0 increases Z_s becomes less and less dependent on η and *vice versa*. Thus for large values of ξ_0 the surface impedance is a constant

$$Z_s \to c.\cosh\xi_0\frac{u^2}{i\omega\kappa_0}\frac{Me_0^{(1)}(\xi_0,-q)}{\frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi,-q)]_{\xi=\xi_0}} \tag{36}$$

while for $\xi \to 0$

$$Z_s \to c.|\sin\eta|\frac{u^2}{i\omega\kappa_0}\frac{Me_0^{(1)}(\xi_0,-q)}{\frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi,-q)]_{\xi=\xi_0}} \tag{37}$$

The wave impedance which is defined by $Z_0 = E_t/H_t$ (the subscript t indicates the transverse component of the field vectors) is a constant

$$Z_0 = (\beta/k)\sqrt{(\mu_0/\kappa_0)} \tag{38}$$

and, of course, is the same as the wave impedance of a circular guide.

Since the space metrical coefficient occurs over and over again in the expression for Z_s it will be convenient to put

$$\overline{Z}_s = Z_s/h \tag{39}$$

The surface impedance, $\overline{Z}_s$, so defined is independent of η and is a characteristic of the guide.

When the eccentricity of the ellipse tends to zero the elliptical guide passes gradually into a circle and c and $q \to 0$, while $\eta \to \phi$ and $ce_0(\eta, -q) \to 1/\sqrt{2}$ thus

$$(\partial/\partial\eta)[ce_0(\eta,-q)] \to 0, \text{ while } \cosh\xi_0 \to \infty$$

but so that $c.\cosh\xi_0 \to r_0$. In consequence H_ξ and $E_\eta \to 0$ while $Me_0^{(1)}(\xi, -q)$ becomes a constant multiple of $H_0^{(1)}(iur)$. Hence the wave becomes the surface wave of the circular cylinder. We can thus conclude that the surface wave of the circular cylinder is stable with respect to a small deformation of the guide cross-section.

The presence of H_ξ and E_η is an inherent consequence of the ellipticity of the guide. The lines of magnetic force do not run along the η co-ordinates, neither need they form closed loops around the guide; on the contrary, some of the lines must penetrate the guide surface as shown in Fig. 8.

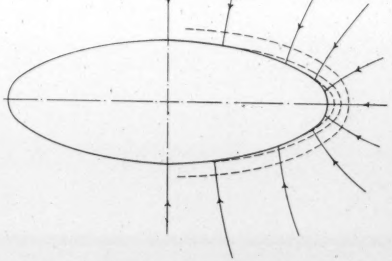


Fig. 8. Lines of force of the elliptic surface wave
——— electric lines; - - - - magnetic lines.

4. SOME PROPERTIES OF SLIGHTLY DEFORMED CIRCULAR GUIDES

A slightly deformed circular guide can be analysed as an elliptical guide of very small eccentricity. An ellipse of small eccentricity is characterized by a large value of ξ on its surface and at the same time a small value of c. For this case, by virtue of equation (19), q will be a very small number.

Now the coefficients A_{2m} in equations (22), (23) and (24) are functions of q and can be expanded in a series[9] of ascending power of q. Thus for small values of q we can neglect all terms of the series except the first two, and in what follows we shall take:

$$\left.\begin{aligned} A_0 &= 1/\sqrt{2} \\ A_2 &= -\,q_0/2\sqrt{2} \end{aligned}\right\} \tag{40}$$

and neglect all the remaining coefficients, so that

$$\left.\begin{aligned} ce_0(\eta, -q) &\simeq \frac{1}{\sqrt{2}}\left(1 + \frac{q}{2}\cos 2\eta\right) \\ \frac{\partial}{\partial\eta}[ce_0(\eta, -q)] &\simeq -\frac{q}{\sqrt{2}}\sin 2\eta \end{aligned}\right\} \tag{41}$$

The derivative of ce_0 is shown plotted in Fig. 9.

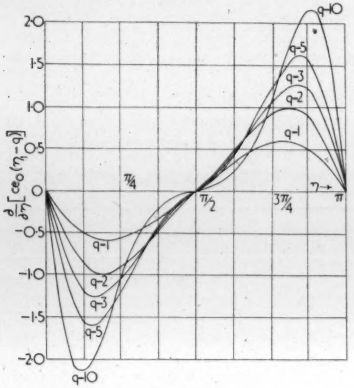


Fig. 9. The function of $\frac{\partial}{\partial\eta}[ce_0(\eta, -q)]$

Further, in the proximity of the point, v_1 and v_2 will be small so that we can replace the Bessel functions in equation (24) by their small argument representation, namely:

$$\left.\begin{aligned} J_0(x) &\simeq 1 \\ J_1(x) &\simeq x/2 \\ H_0^{(1)}(x) &\simeq i(2/\pi)\ln(-i.0{\cdot}89x) \\ H_1^{(1)}(x) &\simeq -\,i(2/\pi).1/x \end{aligned}\right\} \tag{42}$$

so that

$$Me_0^{(1)}(\xi, q_0) \simeq i\frac{\sqrt{2}}{\pi}\left\{\left[1 - \frac{q_0}{4}\exp(-2\xi)\right]\right.$$
$$\left.\times\ln[-i0{\cdot}89\sqrt{q}\exp(\xi)] - \frac{q_0}{4}\exp(-2\xi)\right\} \tag{43a}$$
$$\simeq i\frac{\sqrt{2}}{\pi}\ln[-i0{\cdot}89\sqrt{q}\exp(\xi)] \tag{43b}$$

and

$$\frac{\partial}{\partial\xi}[Me_0^{(1)}(\xi, q_0)] \simeq i\frac{\sqrt{2}}{\pi}\left\{1 + \frac{q_0}{4} - \frac{q_0}{4}\exp(-2\xi)\right.$$
$$\left.\times\ln[-i0{\cdot}89\sqrt{q}\exp(\xi)]\right\} \tag{44a}$$
$$\simeq i\frac{\sqrt{2}}{\pi} \tag{44b}$$

Equations (41), (43*b*) and (44*b*) when substituted in (21) and (32) reveal the nature of the field in close proximity to an almost circular guide. Thus:

$$\left.\begin{aligned} E_z &= i\frac{A}{\pi}.\ln(0{\cdot}89ur).\left(1 + \frac{q}{2}\cos 2\eta\right) \\ H_\xi &= -\frac{A}{\pi}\frac{\omega\kappa_0}{u^2}\cdot\frac{1}{h}\cdot\ln(0{\cdot}89ur)\cdot q\sin 2\eta \\ H_\eta &= -\frac{A}{\pi}\frac{\omega\kappa_0}{u^2}\frac{1}{h}\left(1 + \frac{q}{2}\cos 2\eta\right) \\ E_\xi &= -\frac{A}{\pi}\frac{\beta}{u^2}\frac{1}{h}\left(1 + \frac{q}{2}\cos 2\eta\right) \\ E_\eta &= \frac{A}{\pi}\frac{\beta}{u^2}\frac{1}{h}\ln(0{\cdot}89ur).q\sin 2\eta \end{aligned}\right\} \tag{45}$$

To the same degree of approximation

$$\overline{Z}_s \simeq (u^2/i\omega\kappa_0)\ln(0{\cdot}89ur_0) \tag{46}$$

where r_0 = radius of the guide.

To obtain Z_s to a higher order of approximation equations (41*a*) and (42*a*) must be used. We then obtain

$$Z_s = \frac{u^2}{i\omega\kappa_0}\left[1 - \frac{1}{8}\left(\frac{uc}{2}\right)^2\left(\frac{c}{r}\right)^2\ln(0{\cdot}89ur_0)\right] \times \ln(0{\cdot}89ur_0) \tag{47}$$

This expression for Z_s differs from the corresponding expression for a circular guide[2] by the factor inside the square brackets. The quantity $\frac{1}{8}\left(\frac{uc}{2}\right)^2\left(\frac{c}{r_0}\right)^2\ln(0{\cdot}89ur_0)$ is for all cases to which the formula applies very small (see also example in Section 6.1) and we conclude that for a given value of Z_s a slight deformation of a circular guide has a negligible effect on the value of the radial propagation coefficient.

Further, since the guide attenuation, α, is given by

$$\alpha \simeq \frac{ab}{k} = \frac{|u|^2}{2k}\sin 2\psi \tag{48}$$

where $u = a - ib = |u|\angle\psi$

it transpires that—since a small deformation of the guide geometry brings about but a negligible change in u—the attenuation of a circular surface wave guide is substantially unaffected by a slight deformation of the guide cross-section.

5. THE RIBBON GUIDE

When $\xi \to 0$ the ellipse collapses to a flat strip (Fig. 2) and $\xi = 0$ we take as the definition of a ribbon guide, where $2c$ is the width of the guide.

The field equations for this case are given by equation (32), while the expression for $\overline{Z}_s$ simplifies notably if we restrict our analysis to narrow ribbons only. Thus, when c is small we get [from equations (33), (43*a*) and (44*a*)]

$$\overline{Z}_s = \frac{u^2}{i\omega\kappa_0}\ln\left[.0{\cdot}89u\left(\frac{c}{2}\right)\right]\left\{1 - \frac{1}{2}u^2\left(\frac{c}{2}\right)^2 \times \ln\left[0{\cdot}89u\left(\frac{c}{2}\right)\right]\right\} \tag{49}$$

We thus arrive at the conclusion that a ribbon guide of a given value of $\overline{Z}_s$ has almost the same radial propagation coefficient as a circular guide of the same $\overline{Z}_s$ but of a diameter equal to half the width of the ribbon.

6. EXAMPLES OF ELLIPTIC SURFACE WAVE GUIDES FOR SUPPORT OF PRINCIPAL E-MODES

To obtain an elliptic surface wave we must have a guide, the surface impedance of which is equal to equation (33) at all points on the guide surface. Some of the guides that possess this property are analysed below.

6.1 *Dielectric coated perfectly conducting elliptical rod.* Consider the guide structure illustrated in Fig. 10. The solution to the wave equation inside the dielectric shell is [see also equations (10) and (12)]

$$E_z = [B.Ce_0(\xi, q_1) + C.Fey_0(\xi, q_1)]ce_0(\eta, q_1) \qquad (50)$$

where $$q_1 = (k_1 c/2)^2$$

and $$k_1^2 + \beta^2 = \omega^2 \mu_0 \kappa_1 \qquad (51)$$

while μ_0, κ_1 are respectively the permeability and permittivity of the dielectric coating.

Fig. 10. The dielectric coated elliptical guide

Let ξ_1 and ξ_0 be the co-ordinates of the perfectly conducting inner core and the surface of the guide respectively. Then for the field components inside the coating we obtain

$$\left.\begin{aligned}
H_\xi &= \frac{i\omega\kappa_1}{k_1^2}\cdot\frac{1}{h}\cdot[B.Ce_0(\xi,q_1) + C.Fey_0(\xi,q_1)]\,.\\
&\qquad\qquad \frac{\partial}{\partial\eta}[ce_0(\eta,q_1)]\\
H_\eta &= -\frac{i\omega\kappa_1}{k_1^2}\cdot\frac{1}{h}\cdot\frac{\partial}{\partial\xi}[B.Ce_0(\xi,q_1) + C.Fey_0(\xi,q_1)]\,.\\
&\qquad\qquad ce_0(\eta,q_1)\\
E_\xi &= -\frac{i\beta}{k_1^2}\cdot\frac{1}{h}\cdot\frac{\partial}{\partial\xi}[B.Ce_0(\xi,q_1) + C.Fey(\xi,q_1)]\,.\\
&\qquad\qquad ce_0(\eta,q_1)\\
E_\eta &= -\frac{i\beta}{k_1^2}\cdot\frac{1}{h}\cdot[B\cdot Ce_0(\xi,q_1) + C.Fey(\xi,q_1)]\\
&\qquad\qquad \frac{\partial}{\partial\eta}[ce_0(\eta,q_1)]
\end{aligned}\right\} \qquad (52)$$

where $$\frac{B}{C} = -\frac{Fey_0(\xi_1,q_1)}{Ce_0(\xi_1,q_1)} \qquad (53)$$

And the surface impedance $\overline{Z}_s$ is given by:

$$\overline{Z}_s = -\frac{k_1^2}{i\omega\kappa_1}\cdot\frac{B.Ce_0(\xi_0,q_1) + C.Fey_0(\xi_0,q_1)}{\dfrac{\partial}{\partial\xi}[B.Ce_0(\xi,q_1) + C.Fey_0(\xi,q_1)]_{\xi=\xi_0}} \qquad (54)$$

On equating (33) and (54) we get:

$$-\left(\frac{k_1}{u}\right)^2\cdot\left(\frac{\kappa_0}{\kappa_1}\right)\cdot\frac{B.Ce_0(\xi_0,q_1) + C.Fey_0(\xi_0,q_1)}{\dfrac{\partial}{\partial\xi}[B.Ce_0(\xi,q_1) + C.Fey.(\xi,q_1)]_{\xi=\xi_0}} = \frac{Me_0^{(1)}(\xi_0,-q)}{\dfrac{\partial}{\partial\xi}[Me_0^{(1)}(\xi,-q)]_{\xi=\xi_0}} \qquad (55)$$

The transcendental equation (55) could be solved in principle for u or β. Thus we can conclude that a guide as illustrated in Fig. 10 can support an elliptic surface wave.

The exact solution to equation (55) is impracticable but we can make some progress by approximations. We assume that the coating of the dielectric is very thin so that equation (53) can be approximated to by the first two terms of the Taylor's expansion for the respective functions:

$$\frac{B}{C} \simeq -\frac{Fey_0(\xi_0,q_1) - \Delta\xi.F'ey_0(\xi_0,q_1)}{Ce_0(\xi_0,q_1) - \Delta\xi.C'e_0(\xi_0,q_1)} \qquad (56)$$

where $\Delta\xi = \xi_0 - \xi_1$ and accents over the functions indicate differentiation with respect to ξ.

When equation (56) is substituted in equation (54) after some straightforward algebra we obtain:

$$\left.\begin{aligned}
&\overline{Z}_s \simeq i\omega\mu_0\left(1 - \frac{\kappa_0}{\kappa_1}\right).\Delta\xi\\
\text{or}\quad &Z_s \simeq i\frac{2\pi}{\lambda_0}\sqrt{\left(\frac{\mu_0}{\kappa_0}\right)}.l_\xi.\left(1 - \frac{\kappa_0}{\kappa_1}\right)
\end{aligned}\right\} \qquad (57)$$

where $l_\xi = h.\Delta\xi$ = thickness of coating.

The expression (57) is sufficiently accurate for many practical applications. As an example, let us consider a circular perfectly conducting wire (radius r) coated with a thin layer of dielectric that is not uniform in thickness but varies slightly due to mechanical inaccuracies.

Suppose that the thickness of the dielectric varies according to

$$l_\xi = t(1 - \epsilon.\cos 2\phi)$$

Further, take $t/\lambda_0 = 0{\cdot}002$, $\kappa_0/\kappa_1 = 1/2$, $\epsilon = 0{\cdot}1$, $r_0 = 1$ mm and $\lambda_0 = 3{\cdot}2$ cm.

Now, had $\epsilon = 0$ then we would have

$$X_s = 2{\cdot}37\ \Omega$$

and $$u = 12{\cdot}8\ m^{-1}$$

But since $\epsilon = 0{\cdot}1$

therefore $\overline{X_s} = 2{\cdot}37\ \Omega$

and $X_s = 2{\cdot}37\ (1 - 0{\cdot}1\cos 2\phi)$

However, since the ellipse does not differ appreciably from a circle we can take [from equation (15)]

$$h \simeq r_0\left[1 - \frac{1}{2}\left(\frac{c}{r}\right)^2\cos 2\phi\right] \qquad (58)$$

In consequence, we find that one-half the interfocal distance of the virtual ellipse c, becomes approximately equal to $0{\cdot}45.r_0$. In addition we find that the factor in the square brackets of equation (47) is $1{\cdot}005$. Thus even a 10% distortion in the dielectric coating of a circular guide has a negligible effect on the value of the radial propagation coefficient.

6.2 *The corrugated elliptical guide.* Consider a corrugated elliptical guide fulfilling the following conditions:

(i) the corrugations are rectangular;
(ii) the corrugation pitch is much smaller than the guide wavelength;
(iii) the depth of corrugation is much greater than the width of corrugation;
(iv) the depth of corrugation is small in comparison with the radius of curvature of the guide.

The mathematical analysis for an elliptical corrugated guide is very similar to that for a circular guide, and since the latter has been a subject of a separate paper,[6] there is no purpose in repeating the same. The only difference in the two cases is that, whenever a Bessel function occurs in the analysis for a circular guide, a corresponding Mathieu function is used in the case of an elliptical guide. Further, the guide characteristics (i)–(iv) permit a number of approximations[6] to be used successfully and a simple expression for $\overline{Z}_s$ of a corrugated elliptical guide is easily obtained, which naturally is the same as $\overline{Z}_s$ of a circular guide. Thus, using the result for a circular guide[6] we get for an elliptical guide

$$\overline{Z}_s \simeq i\frac{d}{D}\cdot\frac{2\pi}{\lambda_0}\cdot l_\xi\cdot\sqrt{\left(\frac{\mu_0}{\kappa_0}\right)} \tag{59}$$

where $l_\xi = h\,.\,\Delta\xi$ equals depth of corrugation and

$$\frac{d}{D} = \frac{\text{width of groove}}{\text{pitch of corrugation}}$$

6.3 *Homogeneous metal guide.* Suppose the elliptical guide is formed by a homogeneous solid metal (conductance equals g) rod, then the solution is derived from [see equation (12)]

$$E_z = D\,.\,Ce_0(\xi, q_2)\,.\,ce_0(\eta, q_2) \tag{60}$$

$$\left.\begin{aligned} &\text{Where} \quad q_2 = \left(\frac{k_2 c}{2}\right)^2 \simeq i\frac{\omega\mu_0 g}{4}c^2 \\ &\text{and} \quad k_2^2 = -\beta^2 - i\omega\mu_0 g \simeq -i\omega\mu_0 g \end{aligned}\right\} \tag{61}$$

The above approximations are justified since the conductance g of metals is extremely large, and we are interested in frequencies of the order of 10^9 c/s.

The remaining field components follow from equation (14) on substitution of equation (60)

$$\left.\begin{aligned} H_\xi &= \frac{D}{h}\cdot\frac{g}{k_2^2}\cdot Ce_0(\xi, q_2)\cdot\frac{\partial}{\partial\eta}[ce_0(\eta, q_2)] \\ H_\eta &= -\frac{D}{h}\cdot\frac{g}{k_2^2}\frac{\partial}{\partial\xi}[Ce_0(\xi, q_2)]\,.\,ce_0(\eta, q_2) \\ E_\xi &= -\frac{D}{h}\cdot\frac{i\beta}{k_2^2}\frac{\partial}{\partial\xi}[Ce_0(\xi, q_2)]\,.\,ce_0(\eta, q_2) \\ E_\eta &= -\frac{D}{h}\cdot\frac{i\beta}{k_2^2}Ce_0(\xi, q_2)\cdot\partial/\partial\eta[ce_0(\eta, q_2)] \end{aligned}\right\} \tag{62}$$

$$\text{Thus} \quad \overline{Z}_s = \left.\frac{E_z}{H_\eta}\right|_{\xi=\xi_0} = -\frac{k_2^2}{g}\cdot\frac{Ce_0(\xi_0, q_2)}{\dfrac{\partial}{\partial\xi}[Ce_0(\xi, q_2)]_{\xi=\xi_0}} \tag{63}$$

This expression for the impedance can be approximated as follows. Firstly, the function Ce_0 can be expanded in a Bessel function product series.[4] Thus

$$\left.\begin{aligned} &Ce_0(\xi, q_2) = \frac{p_0}{A_0}\sum_{r=0}^{\infty}(-1)^r A_{2r} J_r(v_1) J_r(v_2) \\ &\text{where} \quad v_1 = \sqrt{q_2}\exp(-\xi),\ v_2 = \sqrt{q_2}\exp(\xi) \\ &\text{and}^{[4]} \quad p_0 = \frac{ce_0(0, q_2)\,ce_0\left(\dfrac{\pi}{2}, q_2\right)}{A_0} \end{aligned}\right\} \tag{64}$$

Secondly, since q_2 is very large, we can use the asymptotic expression for Bessel functions

$$J_r(v) \simeq \sqrt{\left(\frac{2}{\pi v}\right)}\cos\left(v - \frac{r\pi}{2} - \frac{\pi}{4}\right) \tag{65}$$

When equations (64) and (65) are substituted in equation (63) then, after some algebraic transformations, we finally obtain

$$\overline{Z}_s = \frac{1}{c\sinh\xi}(1+i)\sqrt{\left(\frac{\pi f\mu_0}{g}\right)} \tag{66}$$

Consider a slightly deformed circular guide. In this case equations (66) and (46) can be equated and we get

$$R_s(1+i) = r_m\frac{u^2}{i\omega\kappa_0}\ln(0{\cdot}98ur_m) \tag{67}$$

where $$R_s(1+i) = \sqrt{\left(\frac{\pi f\mu_0}{g}\right)}(1+i)$$

equals intrinsic impedance of a good conductor (68)

and $$r_m = c\,.\sinh\xi$$

equals semi-minor axis of the ellipse.

It transpires from equation (67), by comparison with the corresponding expression for a circular guide, that the decay coefficient of a slightly elliptical guide made of a homogeneous conductor is substantially the same as that of the circular guide, the radius of which equals the semi-minor axis of the ellipse.

6.4 *The elliptical dielectric rod waveguide.* Suppose that the guide is made of a homogeneous dielectric the permeability and permittivity of which are μ_0, κ_1 respectively. The field inside the rod for this case is derived from

$$E_z = B\,.\,Ce_0(\xi, q_1)\,.\,ce_0(\eta, q_1) \tag{69}$$

$$\left.\begin{aligned} &\text{where} \quad q_1 = (k_1 c/2)^2 \\ &\text{and} \quad k_1^2 + \beta^2 = \omega^2\mu_0\kappa_0 \end{aligned}\right\} \tag{70}$$

The remaining field components are the same as given by equation (52) but with $C =$ zero. From these equations we find that the surface impedance is given by:

$$\overline{Z}_s = -\frac{k_1^2}{i\omega\kappa_1}\frac{Ce_0(\xi_0, q_1)}{\dfrac{\partial}{\partial\xi}[Ce_0(\xi_0, q_1)]_{\xi=\xi_0}} \tag{71}$$

In the case of slightly elliptical guides equation (71) can be approximated by using only the first two terms of the series expansion for $Ce_0(\xi, q)$. We then obtain

$$Z_s \simeq h\frac{k_1^2}{i\omega\kappa_1}\left[\frac{A_0 J_0(k_1 r) - A_2 J_2(k_1 r)}{(A_0 + A_2)J_1(k_1 r) - A_2\cdot\dfrac{2}{r}\cdot J_2(k_1 r)}\right] \tag{72}$$

Unless the fraction inside the square brackets is a negative quantity there will be no surface wave, and the "cut-off" condition is determined from the equation:

$$A_0 J_0(k_1 r) = A_2 J_2(k_1 r) \tag{73}$$

From this it is evident that the cut-off frequency is slightly higher for the elliptical than it is for the circular guide, the radius of which is equal to the semi-major axis of the ellipse.

7. CONCLUSIONS

An analysis for the surface wave of the elliptic cylinder was carried out and the nature of the wave was discussed and contrasted with the surface wave of the circular cylinder.

It was shown that although the field guided by an elliptical guide decays exponentially at large distances from the axis of the guide, it does not degenerate into the surface wave of the circular cylinder at any distance from the guide. Yet, an elliptic surface wave does gradually pass into a surface of the circular cylinder when the elliptical guide is gradually made to take up a circular shape.

The theory of the elliptic surface wave is of importance when analysing deformed circular guides (e.g. due to mechanical inaccuracies) and in this connexion it was shown that a small deformation (say $\pm 10\%$) of a circular guide has a negligible effect on its performance.

ACKNOWLEDGEMENTS

The author wishes to express his thanks to Professor H. M. Barlow and Dr. A. L. Cullen, of University College, London, for their helpful suggestions during the preparation of the paper.

REFERENCES

(1) GOUBAU, G. *J. Appl. Phys.*, **21**, p. 1119 (1950).

(2) BARLOW, H. M., and KARBOWIAK, A. E. *Proc. Instn Elect. Engrs*, **100**, III, p. 321 (1953).

(3) SOMMERFELD, A. *Ann. Phys. Chemie*, **67**, p. 233 (1899).

(4) HARMS, F. *Ann. Phys.* [*Leipzig*], **23**, p. 44 (1907).

(5) HONDROS, D., and DEBYE, F. *Ann. Phys.* [*Leipzig*], **32**, p. 465 (1910).

(6) BARLOW, H. M., and KARBOWIAK, A. E. *Proc. Instn Elect. Engrs*, **101**, III, p. 182 (1954).

(7) ZENNECK, J. *Ann. Phys.* [*Leipzig*], **23**, p. 846 (1907).

(8) BARLOW, H. M., and CULLEN, A. L. *Proc. Instn Elect. Engrs*, **100**, III, p. 329 (1953).

(9) MACLACHLAN, N. W. *Theory and application of Mathieu Functions* (Oxford: Clarendon Press, 1947).

(10) INCE, E. L. *Proc. Roy. Soc. Edinburgh*, **52**, p. 355 (1932).

(11) STRATTON, J. A., MORSE, P. M., CHU, L. J., and HUTNER, R. A. *Elliptic cylinder and Spheroidal wave functions* (1941).

The electrical breakdown of pre-stressed air at atmospheric pressure

By D. R. HARDY, M.Sc.(Eng.), Ph.D., Department of Electrical Engineering, University of Manchester, and H. WROE, M.Sc., Metropolitan-Vickers Electrical Co. Ltd., Manchester

[*Paper first received* 23 *February, and in final form* 5 *March*, 1954]

Time lags to breakdown are measured for surge voltages applied to gaps, which are irradiated with radium, and across which exist steady fields, both uniform and non-uniform. An appreciable difference in time lags for a given overvoltage is observed as the ratio of approach to surge voltage is varied.

1. INTRODUCTION

Workers in the field of electrical breakdown in gases who have been concerned with the measurement of time lags have, for convenience, generally used a test voltage consisting of two parts. A steady (d.c.) pre-stress or approach voltage is applied to the gap and then, at the instant from which the time lag to breakdown is to be measured, a surge voltage of a magnitude equal to the difference between the required test voltage and the approach voltage is superimposed.

The approach voltage may take any value from zero up to the d.c. breakdown value so that, as the approach voltage is increased, the surge must be correspondingly reduced in order to maintain a constant test voltage. If the approach voltage is increased sufficiently, ionization by collision will take place and the resulting space charge in the gap might be expected to affect the breakdown characteristics. Evidence to the contrary has, however, been put forward but the problem was not investigated systematically; in most cases results obtained at two values of approach voltage were compared and found to be approximately the same.

Radium is sometimes preferred to ultra-violet radiations as a source of irradiation, since the ionization produced is constant over very long periods and is less dependent on electrode material and surface condition, but because of the random nature of the effective γ-rays, the results may not be compared directly, since photoelectrons produced by ultra-violet irradiation are distributed uniformly in time.

It is the object of this paper to investigate the effect of approach voltage over the complete range for gaps with near uniform and non-uniform fields in which the initiating electrons are produced by radioactive materials.

2. PREVIOUS WORK

Tilles,[1] investigating time lags in uniform fields at low overvoltages, used two waveshapes, one with zero approach voltage and the other with an approach voltage of 96% of the d.c. breakdown value. The time lags to breakdown, although largely random in nature, were found to be longer by about a factor of 2 with zero approach voltage than with the 96% value, when the gap was weakly irradiated with ultra-violet light. It was predicted that, had the approach voltage been lowered, the time lags would have increased until, at a low value, a maximum time lag perhaps much larger than those measured would have been obtained.

White [2] measured time lags, using a Kerr cell technique, for high overvoltages and intense ultra-violet irradiation produced by an auxiliary spark. The test gap was supplied with an approach voltage about equal to the d.c. breakdown value. The fact that the time lag was measured from the instant that the irradiation was applied may introduce serious errors since it has been shown[3] that the time taken for the illumination to attain its maximum value is of the same order of magnitude as the measured time lags. The effect of approach voltage on the breakdown mechanism

may be negligible until the auxiliary spark has been initiated (that is, during the actual time lag period) and could account, in part, for the long lags observed.

This work was continued by Wilson,[4] using ultra-violet irradiation produced by a steady mercury vapour light source. An approach voltage of 99% was applied to one sphere and a pulse of the appropriate amplitude and of opposite polarity was applied to the other by means of a vacuum switch. No oscillographic record was taken of the waveshapes so produced and it has been since suggested[5] that severe oscillations may have occurred at the switch contacts, thus seriously affecting the results. The spheres were cleaned after every fifty sparks (that is, in the middle of a test) and, since no effort was made to measure the actual gap current, but merely the incident radiation, the photocurrent from the cathode may have changed considerably.[6].

Further studies were made by Newman,[7] who used a travelling wave technique to measure the time lags. No difference was observed as the approach voltage was changed from 50 to 90% of the d.c. breakdown value for small gaps between 0·75 in. diameter brass spheres.

An investigation of statistical time lag distributions was made by Cobine and Easton[5] but no approach voltage was used. Some criticism was levelled at the work of Tilles in that the distribution of time lags might be better represented by a modified expression which depends on the time function of the applied test wave.

Fisher and Bederson[8] measured time lags with very low overvoltages at atmospheric pressure using an oscillograph technique. The approach voltages used were not expressed as a percentage of the d.c. breakdown value, but are estimated at 82 and 64% for a 0·3 cm gap. No difference in time lags was observed for these values.

3. EXPERIMENTAL METHODS

In the test gap used for the near uniform field experiments, the electrodes consisted of 2 cm diameter brass spheres and, for the non-uniform fields, of a high voltage point tungsten electrode and an earthed brass plane of 4 cm diameter. The whole gap was screened to exclude any stray ultra-violet light. The upper electrode was mounted on a micrometer head adjustable to 0·002 mm and the lower electrode on an insulating pillar of polystyrene, so that the input resistance to an electrometer amplifier, immediately below the gap, was maintained at a high value. The source of irradiation, used to provide a copious supply of electrons in the gap, was in the form of a 1 mg radium capsule which was contained behind and within a short distance of the sparking point of the upper sphere in the case of the spherical electrodes, and behind the earthed plane in the case of the non-uniform field experiments.

Before each set of results the electrodes were polished with very fine emery paper and then with a clean dry cloth. Care was taken to keep the electrodes smooth and free from pitting and a high polish was avoided.

The ionization produced was estimated from the measured value of gap current which attained a more or less constant value of 9×10^{-12} A, with about 150 V applied to a 0·1 cm gap between 2 cm diameter spheres, thus comparing favourably with the value given by Hardy and Craggs.[6]

Time lags were recorded by means of a high-speed transient oscillograph used in conjunction with a compensated resistance-capacitance type of potential divider. This method, although laborious, enabled the shape of the voltage wave to be observed, undesirable oscillations to be detected and, what is important[9] in the short time lag domain, the wavefront time to be measured. A zero line and the output from a standard oscillator were superimposed on each oscillogram and the tube sensitivities assumed constant for the few seconds which elapsed between the recording of the oscillogram and the calibration wave.

The power supplies were of conventional design and, with the small loads used in the course of the present experiments, ripple was negligible. Since the overvoltage used was never less than 10%, a highly stabilized voltage source was considered unnecessary. The d.c. voltages were measured by means of a microammeter and a number of high stability carbon resistors, and the conditions of operation chosen to give an estimated error of less than 1%.

The intense ultra-violet irradiation from a nearby spark discharge was used to trip a single stage impulse generator as shown in Fig. 1. This method has the advantage of no

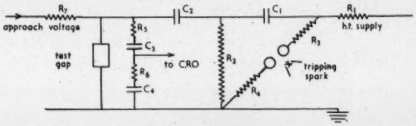


Fig. 1. Circuit for the application of surge voltage to pre-stressed gap

R_1, $R_2 = 10^7\ \Omega$
R_3, $R_4 = 50\ \Omega$
$R_5 = 1100\ \Omega$
$R_6 = 33\ \Omega$
$R_7 = 2 \times 10^8\ \Omega$

$C_1 = 0{\cdot}04\ \mu F$
$C_2 = 500$ pF
$C_3 = 15$ pF
$C_4 = 500$ pF

direct connexion between the tripping pulse generator and the main circuit and thus reduces the likelihood of spurious oscillations in the measuring circuit. Care had to be observed when setting the trigger gap since the working range with this method of tripping is small.

4. RESULTS

To estimate accurately the most probable time lag for a given set of conditions would require a large number of measurements for each test, whereas fewer measurements would reduce accuracy but at the same time enable a wider field to be covered. As a compromise twenty measurements were made for each test, the average value being shown as a point on a line at a particular overvoltage and the range of results indicated by the length of the line. In some cases the lags were too short to measure, in which event the time spread line is shown in dotted form on the appropriate figure.

The terms percentage overvoltage and approach voltage are used throughout the paper and are defined as follows:

$$\text{Approach voltage} = \frac{V_a}{V} \times 100\%.$$

$$\text{Overvoltage} = \frac{V_a + V_s}{V} \times 100\%$$

where V_s = surge voltage;
V_a = approach voltage;
V = d.c. breakdown voltage.

4.1. *Uniform fields*

All experiments were carried out with a 0·1 cm gap between 2 cm diameter spheres so that the field form was approximately uniform in the sparking area. Tests were made with a range of approach voltages from 0 to 97·6% and with overvoltages of from about 10 to 300%.

The results obtained for selected values of approach voltage are shown in Fig. 2. The curve [Fig. 2(*a*)], representing zero approach voltage, shows that a large increase in time lag occurs as the overvoltage is reduced below 50%. The lags obtained with 30–50% overvoltage are found to lie within two domains, a number being too short to measure with the time base at the setting required to record the longer lags.

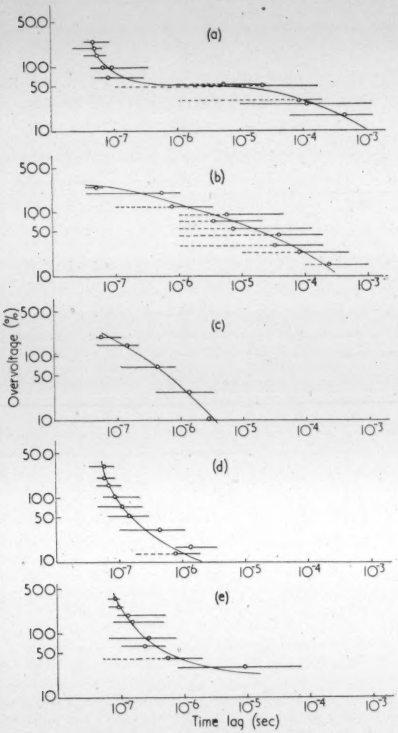


Fig. 2. The effect of overvoltage on time lag to breakdown for various pre-stress voltages. 0·1 cm gap between 2 cm diameter spheres. 1 mg radium

(*a*) Zero pre-stress voltage.
(*b*) Pre-stress voltage 0·1% of d.c. breakdown voltage.
(*c*) Pre-stress voltage 48% of d.c. breakdown voltage.
(*d*) Pre-stress voltage 70% of d.c. breakdown voltage.
(*e*) Pre-stress voltage 97% of d.c. breakdown voltage.

The application of an approach voltage of 4 V, which is less than 0·1% of the d.c. breakdown value, produced a noticeable change in the time lag characteristic, as shown in Fig. 2(*b*). The average lag is increased by a factor of 20 or 30 at 100% overvoltage. With increases in approach voltage to 2 and 4%, the time lags were found to increase but further increases in approach voltage to 10 and 20% produced a reduction in the values of average time lag. The curve obtained for 48% approach voltage is given in Fig. 2(*c*), from which it will be seen that at the lower overvoltages the time lag is reduced by some several hundred times. An approach voltage of 70%, shown in Fig. 2(*d*), gives the lowest values of time lags, and a further increase again produces longer lags. An example of this is given in Fig. 2(*e*) which is for an approach voltage of 97%.

These results are summarized in Fig. 3 in which are plotted the average time lags as a function of approach voltage for overvoltages of 20, 50, 100 and 200%. The scatter of the points in Curve 4 may be attributed to the small gradients of the time lag-overvoltage curves for low approach voltages, since a small displacement of these curves leads to a large

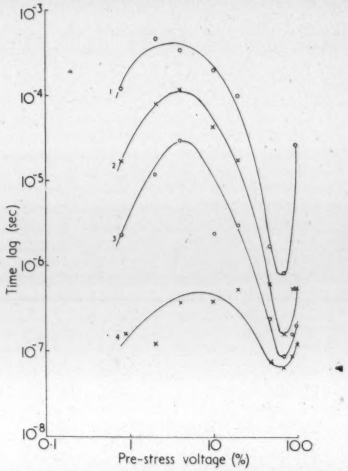


Fig. 3. The effect of pre-stress voltage on time lag to breakdown for various overvoltages (both voltages expressed as a percentage of the d.c. breakdown value). 0·1 cm gap between 2 cm diameter spheres. 1 mg radium

Curve 1, 20% overvoltage.
Curve 2, 50% overvoltage.
Curve 3, 100% overvoltage.
Curve 4, 200% overvoltage.

change in time lag. At the lower approach voltages the time lag tends to increase as the approach voltage increases. This may be ascribed to the removal from the gap of the ionized particles produced initially by the radium capsule. Indeed the gap current was observed to approach a more or less constant value at a voltage of about 150 V, which corresponds to the approach voltage of about 3% at which the maximum measured time lags occurred.

It is unlikely that the reduction in time lag, which occurs for approach voltages greater than about 4%, is due to the decreasing amplitude of the surge voltage, since the only factor which may affect the lag is the time of the wavefront and this is a negligible fraction of the average time lag in these experiments.

The increase in approach voltage, however, will ultimately produce values of X/p at which α and γ mechanisms become

active. α increases rapidly in air, at atmospheric pressure, at a value of $X/p = 31$[10] which, in the present work, corresponds to an approach voltage of 50%. Again, the γ mechanism in air has been observed[11] at $X/p = 39$ so that an approach of about 60% is required for this mechanism to be active.

However, the reduction in time lag as the approach voltage is increased above 4% (which corresponds to $X/p = 3$ approximately) may be accounted for by local field intensification at the cathode due to irregularities caused by pitting of the spheres and, possibly, by oxidation.[11] An increase in field of tenfold at least is therefore required to produce an effective secondary mechanism. Such enhanced fields have, in fact, been observed. Schottky,[12] investigating field emission at low pressures, estimated that for submicroscopic irregularities of 10^{-5}–10^{-6} cm radius, a value for the ratio of local to average field at the cathode of 10 is not unusual. More recently, Llewellyn-Jones[13] has measured surface projections of about 10^{-3} cm on electrodes between which many sparks have passed.

As the approach voltage is increased above 4%, the local field, and therefore the production of electrons by secondary processes, will be enhanced, thus increasing the probability of breakdown when the surge voltage is applied and reducing the statistical time lag. At an approach voltage of about 50% the α mechanism becomes noticeably active and a positive ion space charge builds up in the gap. A further increase to 60% produces a tenfold increase in α and the resulting space charge causes appreciable field distortion which ultimately limits further development. The increased time lags in this region suggest that the positive space charge inhibits the progress of further electron avalanches produced when the surge voltage is applied. Similar effects have been explained[14] in terms of streamer propagation, conditions for which are more favourable as the overvoltage is increased.

4.2. *Non-uniform fields*

No appreciable corona was observed before breakdown with a gap of 0·1 cm between a high-voltage point electrode and an earthed plane. Since it is to be expected that corona intensified space charge would have an appreciable effect on the breakdown of the pre-stressed gap, the spacing was increased to 0·2 cm so that the magnitude of this effect might be investigated.

4.2.1. *Negative point electrode.* Corona began to appear, with a high-voltage cathode, at an approach voltage of 75% and the familiar "mushroom" discharge was seen at an approach voltage of about 84%.

The time lag curve obtained with zero approach voltage, shown in Fig. 4(*a*), is seen to be very similar to but with a longer time lag for a given overvoltage than the curve for the uniform field [Fig. 2(*a*)]. With very slight reductions in time lag, this same type of curve was followed as the approach voltage was increased to 71%. A further increase to 84% produced a pronounced reduction in time lag and at 95%, the results for which are shown in Fig. 4(*b*), a further slight decrease was observed.

The time lag-approach voltage characteristic for the non-uniform gap changes noticeably as the overvoltage is increased. The results for overvoltages of 100 and 60% are summarized in Fig. 6 [Curves 2(*a*) and 2(*b*) respectively]. Although 1 mg of radium was used in these experiments the irradiation produced had but little effect because, presumably, of the small effective gap volume. When corona appeared, the ionization produced was several orders greater than that due to the radium.

With a high-voltage point cathode the initial increase in gap current is not as effective as with a uniform field. The cathode area is small and the field in this region very much higher than in the vicinity of the anode so that, as the gap voltage is increased, ionization by collision occurs more readily. As the approach voltage is increased to 10% a comparatively small increase in average time lag occurs when the overvoltage is 60%, but at the higher overvoltage the lag decreases.

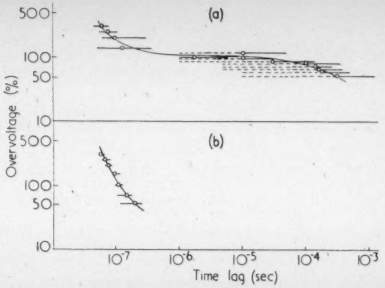


Fig. 4. The effect of overvoltage on time lag to breakdown for various pre-stress voltages. 0·2 cm negative point–plane gap. 1 mg radium

(*a*) Zero pre-stress voltage.
(*b*) Pre-stress voltage 95% of d.c. breakdown voltage.

As the approach voltage is further increased it would appear that the space charge field inhibits the progress of the electron avalanche at the higher overvoltages but not at the lower values. With this electrode configuration it seems reasonable to expect that a high overvoltage would be required in order to produce a space charge of sufficient intensity to prevent the progress of further avalanches. In both cases, however, at an approach voltage of about 70% the time lags decrease rapidly as the approach voltage is increased. A similar effect observed elsewhere[15] was thought to be due to a negative streamer resulting from a negative ion space charge in the vicinity of the anode, since a highly concentrated positive ion space charge would not encourage a cathode mechanism.

4.2.2. *Positive point electrode.* Results taken over a similar range, but with a positive high-voltage point electrode, show an appreciable change in breakdown characteristics from those obtained in the previous section. This is to be expected since no visual corona was observed even with the highest approach voltages.

Time lag curves for zero approach voltage are shown in Fig. 5(*a*), from which it will be seen that the lags are not now in two domains as was the case in the previous sections, but a fairly steady decrease in time lag occurs as the overvoltage is increased. As the approach voltage increases the lags increase slightly and then, at about 10% approach voltage, begin to decrease until at nearly 100% they are again of the same order as those obtained with a few per cent approach voltage. This is illustrated by the curve for 97% approach voltage in Fig. 5(*b*).

The results are summarized, again for overvoltages of 100 and 60%, in Fig. 6 [curves 1(*a*) and 1(*b*) respectively]. The initial scavenging of the gap seems to be more effective in increasing the time lags at the lower approach voltages in both cases. This might be expected since the field at the

cathode is now much less than that in the vicinity of the anode so that the initial current will increase, as the gap voltage is increased, to values in excess of that at which saturation occurs with a uniform field.

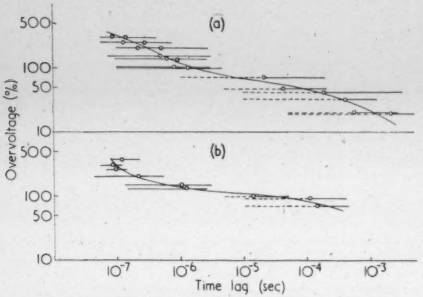


Fig. 5. The effect of overvoltage on time lag to breakdown for various pre-stress voltages. 0·2 cm positive point–plane gap. 1 mg radium

(*a*) Zero pre-stress voltage.
(*b*) Pre-stress voltage 97% of d.c. breakdown voltage.

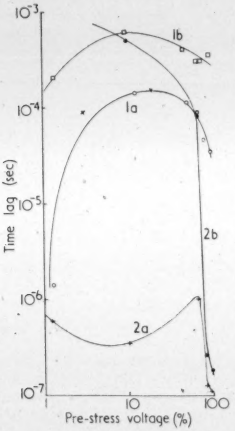


Fig. 6. The effect of pre-stress voltage on time lag to breakdown for various overvoltages (both voltages expressed as a percentage of the d.c. breakdown value). 0·2 cm point–plane gap. 1 mg radium

Curves 1. High-voltage positive point.
Curves 2. High-voltage negative point.
(*a*) 100% overvoltage.
(*b*) 60% overvoltage.

As the approach voltage is still further increased, the secondary processes described in Section 4.1 become active. Because of the increased ionization produced at the cathode the statistical time lags decrease, thus resulting in an increased probability of breakdown. The fact that the time lags were not observed to increase at the higher approach voltages is thought to be due to the geometry of the gap which makes the positive ion space charge more diffuse and, therefore, less effective.

5. CONCLUSIONS

The effect of approach voltage is seen to be appreciable, particularly at the lower overvoltages, with gaps in which the field is nearly uniform. At an overvoltage of 20% the average value of the longest lag, which occurs at an approach voltage of about 3%, is some fifty times that at about 70% approach voltage. At the highest overvoltage investigated (200%) the ratio of longest to shortest lags is reduced to less than 10. Although predicting the general shape of the curves, Tilles did not anticipate the increase in time lag at the highest approach voltages.

With non-uniform fields and a high-voltage point cathode, the effect of corona intensified space charge is most pronounced, and is clearly seen in the curves of Fig. 6. The time lag decreases rapidly, by a factor of over 100 with a few per cent increase in approach voltage, as the onset potential is exceeded. When the overvoltage is high, conditions are favourable for the formation of a positive ion space charge which is able to inhibit the progress of further avalanches.

With a positive high-voltage point electrode, the time lag increases most noticeably as the approach voltage is initially increased, but does not decrease apreciably at the higher values of approach voltage.

This electrode arrangement does not encourage the formation of a concentrated positive ion space charge, nor is corona appreciable, so that the appropriate characteristics are not observed in the relevant curves of Fig. 6.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. F. C. Williams for facilities provided and for permission to publish this paper.

One of us (H.W.) gratefully acknowledges financial assistance given by the Department of Scientific and Industrial Research and included the work described as part of a thesis submitted for the M.Sc. degree of the University of Manchester.

REFERENCES

(1) TILLES, A. *Phys. Rev.*, **46,** p. 1015 (1934).
(2) WHITE, H. J. *Phys. Rev.*, **49,** p. 507 (1936).
(3) SNODDY, L. B. *Phys. Rev.*, **40,** p. 409 (1932).
(4) WILSON, R. R. *Phys. Rev.*, **49,** p. 1082 (1936).
(5) COBINE, J. D., and EASTON, E. C. *J. Appl. Phys.*, **14,** p. 321 (1943).
(6) HARDY, D. R., and CRAGGS, J. D. *Trans Amer. Inst. Elect. Engrs*, **69,** p. 584 (1950).
(7) NEWMAN, M. *Phys. Rev.*, **52,** p. 652 (1937).
(8) FISHER, L. H., and BEDERSON, B. *Phys. Rev.*, **81,** p. 109 (1951).
(9) HARDY, D. R., and LATHAM, H. To be published.
(10) SANDERS, F. H. *Phys. Rev.*, **41,** p. 667 (1932); **44,** p. 1020 (1932).
(11) LLEWELLYN-JONES, F., and PARKER, A. B. *Proc. Phys. Soc.* [*London*], *A*, **213,** p. 185 (1952).
(12) SCHOTTKY, W. *Z. Phys.*, **14,** p. 63 (1923).
(13) LLEWELLYN-JONES, F. *Proc. Phys. Soc.* [*London*] *B*, **62,** p. 366 (1949).
(14) LOEB, L. B. *Fundamental Processes of Electrical Discharges in Gases*, p. 427 (New York: J. Wiley and Sons Inc., 1939).
(15) LOEB, L. B. *Fundamental Processes of Electrical Discharges in Gases*, p. 519 (New York: J. Wiley and Sons Inc., 1939).

Correspondence

Reflectance measurements on siliceous dusts and powders

Using the type S.P. 500 spectrophotometer, by Unicam Instruments (Cambridge) Ltd., graphs of optical reflecting power against wavelength were plotted (Fig. 1) for the following dusts: quartz, chalcedony, flint, Lochaline sand, opal, olivine, garnet, and ground fused silica (Vitreosil). Although all the materials used were highly-pure forms of

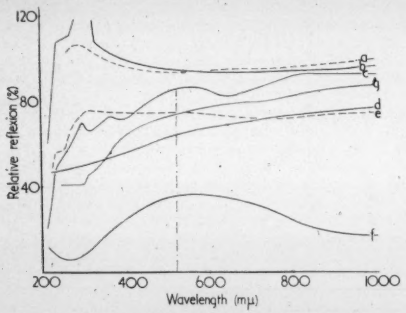


Fig. 1. Characteristic reflectance curves for different siliceous materials in powder form

(*a*) 1 μ rock crystal; (*b*) fused silica (Vitreosil) sieved to 250 μ; (*c*) semi-opal sieved to 250 μ; (*d*) flint sieved to 250 μ; (*e*) chalcedony sieved to 250 μ; (*f*) olivine sieved between 177 and 250 μ; (*g*) Lochaline sand sieved between 177 and 250 μ.

silica (all over 90% silicon dioxide) they showed well-defined differences, and comparison of the graphs for olivine and semi-opal with that of a typical green pigment show that all three have the same dominant wavelength, whilst the Lochaline sand shows many of the characteristics of a yellow pigment.

Characteristic curves were plotted for powders within the same sieve range (177–250 μ) wherever possible, since it was found that, although the reflectance was in general increased by reducing particle size, the salient features of the graphs tended to become less pronounced.

The reflectance depends upon whether the powders have been freshly ground or have been stored. In view of the importance attached by some schools to the "freshly fractured surface theory of silicosis" this is an extremely interesting experimental observation. Thus Lochaline sand which had been stored for many months was taken and the 0–149 μ fraction was sieved out. The larger particles were then ground until they too passed through the 149 μ sieve and they showed an overall increase in percentage reflectance of about 10% above the stored material. (A closer range of particle size was desirable but not immediately possible.)

Ritchie and his co-workers [1,2,3] brought forward evidence to support the theory that a layer of amorphous silica lay on the surface of quartz particles, and compared it with the Beilby layer found on polished surfaces. Several estimates of its thickness using different methods of measurement were given. Their methods of extractive solubility were laborious, and finally, using the electron microscope, they modified their postulate to state that the amorphous layer itself lay on a transition layer of crystallites above the central core. They estimated the thickness of the amorphous layer to be 0·03–0·06 μ.

The reflectance method gave a simple and direct way of confirming the postulate; the 1 μ rock crystal dust was treated with 40% analar hydrofluoric acid for various periods of time and a graph of the type shown in Fig. 2 was obtained.

This shows that the reflectance first falls rapidly then increases to within 1% of its initial value until it finally achieves a steady value corresponding to the homogeneous core. Thus the reflectance measurements confirm the layer theory, but suggest the existence of a greater number of discrete layers

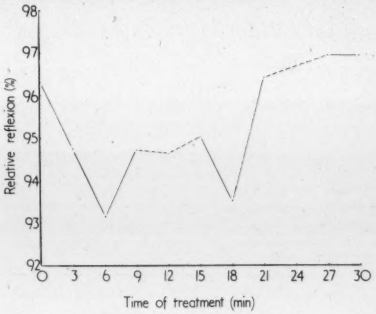


Fig. 2. Reflectance of 1 μ rock crystal after treatment with 40% hydrofluoric acid for various periods $\lambda = 500$ mμ

than was at first thought. It certainly indicates that the gradual and regular change within a single layer is open to doubt.

It may be argued that the change in particle size can account for the change in reflectance. However, using Dale and King's[4] estimate that hydrofluoric acid eats away half the silica by weight in thirty minutes, the change in particle diameter after this time would be only 0·2 μ, giving approximately 0·04 μ change for six minutes of reaction time. It seems doubtful whether this change alone could account for the 3% change in reflectance which is noted, and in any case (from previous graphs) the general inference seems to be that the decrease in particle size would mean an increase in reflectance and not the decrease which is at first observed on the experimental graph.

It is tentatively suggested, therefore, that the particles consist of a homogeneous core surrounded by several discrete layers which themselves, in all probability, show a steady change in properties throughout their individual extent.

From these results, and using the estimate that half the silica is dissolved away after thirty minutes treatment in 40% hydrofluoric acid, the following preliminary estimates of the thickness of each layer may be given, assuming the particles to approximate to spheres:

outer layer	$\simeq$ 0·02 μ;
three inner layers	$\simeq$ 0·03–0·04 μ;
total thickness of layers before reaching core	$\simeq$ 0·12 μ.

This latter figure is in agreement with the overall figure given by Dempster and Ritchie[2] from differential thermal analysis, whilst the first figure compares favourably with the estimate of the amorphous layer thickness given by electron diffraction.

Thanks are due to Mr. J. A. Waddams for the facilities of his laboratories and to Mr. G. Rushton, without whose excellent workmanship in constructing the specimen holder these experiments would not have been possible.

Wolverhampton and Staffordshire Technical College. E. V. SMITH

REFERENCES

(1) CLELLAND, D. W., CUMMING, W. M., and RITCHIE, P. D. *J. Appl. Chem.*, **2**, pp. 31–41 (1952).
(2) DEMPSTER, P. B., and RITCHIE, P. D. *J. Appl. Chem.*, **3**, pp. 182–92 (1953).
(3) GIBB, J. G., RITCHIE, P. D., and SHARPE, J. W. *J. Appl. Chem.*, **3**, pp. 213–18 (1953).
(4) DALE, J. C., and KING, E. J. *American Medical Assoc., Archives of Industrial Hygiene and Occupational Medicine*, **7**, p. 484 (1953).

Application of a tracer to cathode-gettering and gas-adsorption problems

The residual gas pressure of cathode-ray tubes, especially of tubes with a rather small volume such as oscilloscope tubes, decreases when the cathode is heated, even if the anode current is cut off. This effect is an indication that the barium-strontium oxide layer of the cathode might absorb part of the residual gas. It is difficult to check this hypothetical gettering effect of the oxide layer by conventional experimental methods and so we thought of applying a gaseous tracer. The most interesting gas to try in this absorption test is carbon dioxide, as it may poison the oxide layer. It is possible to prepare carbon dioxide with the radioactive ^{14}C-isotope. ^{14}C emits a weak homogeneous β-radiation with an energy of 0·156 MeV; this means that the β-rays are practically absorbed by the glass used for cathode-ray tubes, which has a thickness of 1·0–2·0 mm. The half-life of ^{14}C is 6400 $\pm$ 200 years, thus ensuring no appreciable decline in the total radiation energy of the tracer applied during the tests. ^{14}C can be obtained by the reaction $^{14}N(n, p)$ without any additional radioactive products, which could complicate the measurements.

We used 1 μc ^{14}C in the compound $BaCO_3$, contained in a metal tube, connected to a 5 in. oscilloscope tube. A normal gun with a barium–strontium oxide cathode was mounted in this tube. The components of the gun were made of stainless steel (Fe–Cr–Ni–alloy), but one screening diaphragm consisted of stainless steel, stabilized by titanium. For several reasons we found it interesting to investigate the adsorption of gas by these two materials.

A barium–aluminium flash-getter was mounted in a side tube to assure a low starting pressure (being about 10^{-6} cm of mercury). The tube was assembled and processed in the following way:

Pumping time (total) = 150 min
Temperature of the bulb (maximum) = 380° C (during about 60 min)

Cathode activation:

Cathode temperature ≃ 1000° K during 2 min
Cathode temperature ≃ 1150° K during 2 min
Cathode temperature ≃ 1295° K during 2 min

(The cathode temperature was measured under identical conditions on five fully-activated cathodes by means of a pyrometer and the average values taken into account. Under cathode temperature the black body temperature is to be understood.)

The electrodes were outgassed by h.f. heating (temperature about 800° C during 1–2 min). After outgassing the cathode temperature was 1295° C during 1 min. Then the tube was sealed off, the getter was flashed and the activation of the cathode continued (cathode temperature 1200° K for 30 min, no emission, and at the same cathode temperature emission with 110 V a.c. on grid 1).

Then the residual gas pressure was measured (being about 10^{-6} cm of mercury, as mentioned above) by using the tube itself as ionization chamber. After sealing off the getter tube, the $BaCO_3$-container was heated and $^{14}C^{16}O_2$ produced. Again the gas pressure was measured immediately after the production of carbon dioxide (maximum about 9×10^{-6} cm of mercury); for this measurement an anode current of 400 μA was used during about 20 sec. As the pressure was not stable, due to variations in the rate of carbon dioxide production and immediate absorption of the gas (the cathode being at 300° K all the time and the anode current cut off) it was rather difficult to estimate the actual amount of carbon dioxide introduced. The cathode was heated (temperature 1100° K) for some time (about 30 min) with the anode current cut off.

Thereafter, the $BaCO_3$ container was sealed off, the tube demounted, the cathode and gun components, consisting of the two materials in question, removed and their activities measured by a Geiger–Müller counter. The distance between the samples and the window of the counting tube was about 1 mm.

As the activities of the samples were relatively weak, the counting of each sample was continued for 10 min. This precaution provided a good separation of the activities measured and the background. The measurements of the activities of the samples were repeated ten times and the background was measured after each measurement of the activity of a sample. Then the average values were calculated. The measured Geiger–Müller counts of the samples are listed in Table 1.

Table 1. *Measured counts for cathode, diaphragm and other gun components*

Sample	*Average number of pulses after 10 min counting-time, without background*
Barium–strontium oxide layer* of the cathode	185
Stainless steel, stabilized by titanium	210
Stainless steel	160

* As the surface of the cathode is very porous, part of the activity measured will originate from $^{14}C^{16}O_2$ in the pores. The radiation of these ^{14}C atoms, however, will have been absorbed partly by the oxide layer.

The average number of background pulses after 10 min was 280.

The areas of the surface of the samples varied considerably. To provide relative comparison, the number of pulses after 10 min counting time (without background) divided by the surface in square millimetres (being the "specific activity") was taken as an indication of the amount of gas gettered or adsorbed by the different samples per square millimetre of surface and called the "specific gettering power." Thickness effects were neglected; this means that all the radioactive

material absorbed by the samples were assumed to be located on the surface.

Table 2. *Measured counts per square millimetre for cathode, diaphragm and other gun components*

Sample	*Specific activity (number of pulses after 10 min counting time, per square millimetre of surface, without background)*
Barium-strontium oxide layer* of the cathode	51·00
Stainless steel, stabilized by titanium	0·48
Stainless steel	0·11

* As the surface of the cathode is very porous, part of the activity measured will originate from $^{14}C^{16}O_2$ in the pores. The radiation of these ^{14}C atoms, however, will have been absorbed partly by the oxide layer.

No substantial faults were introduced by this simplification, as the absorption coefficient is rather high for such a weak β-radiation, especially in the case of the stainless steel samples. It was further assumed that the time of exposing the samples to $^{14}C^{16}O_2$ (about 30 min as mentioned above) was sufficient to saturate their gettering ability. [This assumption is not improbable, as after having heated the cathode in a normal 5 in. tube for this length of time (with the anode current cut off) no further drop in gas pressure could be found by prolonged heating of the cathode.] Table 2 shows the specific activities of the three samples tested.

The measurements were repeated with a second tube and the results confirmed.

Comparing the results listed in Table 2, it can be seen that stainless steel, stabilized by titanium, shows about four times the specific gettering power of stainless steel without titanium; the barium-strontium oxide layer of the cathode shows about 500 times the specific gettering power of stainless steel.

Summarizing, it can be said, that the method chosen is suitable to act as a direct proof of the gettering ability of the heated barium–strontium oxide layer of the cathode. Furthermore, it was possible to check on the adsorption of gas by the gun components, this effect being about 500 times weaker than the effects of the cathode layer.

N. V. Philips Gloeilampenfabrieken,
Eindhoven,
Netherlands.

F. DE BOER
W. F. NIKLAS

New books

Introduction to tensors, spinors and relativistic wave equations (relation structure). By E. M. CORSON, Ph.D. (Glasgow: Blackie and Son Ltd., 1953.) Pp. xii + 221. Price 55*s*.

Quantum field theory has now become one of the most important branches of theoretical physics and very many research workers are active in this very sophisticated and often perplexing subject. As always, most of these investigators are far too busy with their research problems to spend time in reviewing the situation—the writing of a book in which the background techniques are assembled and discussed systematically is an even more remote possibility. Dr. Corson's book which sets out to provide this background is therefore especially welcome. It cannot be said that any attempt is made to present the material in a form comprehensible to any but a mathematical physicist who has already had some experience in at least the elements of field theory and relativistic quantum mechanics. Nevertheless, for one who is already working in this very formal branch of modern physics the book should be very valuable. It is certainly systematic and concentrates with the maximum economy of expression on the mathematical techniques which are useful.

Part I is termed "Mathematical introduction." It consists of chapters on tensor analysis and spinor analysis which include at least a cataloguing of the formulae required in Part II. This is entitled "Physical principles," and includes chapters on general field theory, relativistic wave equations—field aspect and relativistic wave equations—matrix algebraic aspect. There is much of interest in these chapters even though they demand a high degree of familiarity with the formal algebra which has been presented in a far from elementary form in Part I. Such important questions as the relation of spin and statistics and the need for second quantization are discussed in a very interesting fashion; so much so that it is a pity that the high degree of formality will severely restrict the audience reached. It is very important to be clear as to what is intrinsic in quantum field theory and what is arbitrary, and this is rendered more difficult at present by the exclusiveness of the subject—the desirability of a physical, rather than a mathematical survey of the subject, as a complement to Dr. Corson's book, cannot be overstressed.

The format of the book is excellent. The printers are to be congratulated on the skill with which they have set up such complicated mathematical formalism in a way which is very easy to read and elegant in appearance. From this point of view the book could not be surpassed.

H. S. W. MASSEY

Elementary introduction to molecular spectra. By B. BAK. (Amsterdam: North-Holland Publishing Co., 1954.) Pp. x + 125. Price 18*s*.

This monograph is intended to provide a brief review of the whole range of molecular spectra for biologists, chemists and chemical engineers who, using spectroscopic methods in solving their problems, wish to broaden their understanding of the subject.

With little mathematics the opening chapter briefly surveys both practical and theoretical aspects of the subject. The second sketches the mathematical derivation of important results of quantum mechanics used in the interpretation of spectra. The remaining three chapters deal with the microwave, the infrared and the visible and ultraviolet regions, respectively; the reader of these chapters needs to have attempted only one short section (about two pages) of the preceding chapter. The references in the text are to larger recent books. A subject index is provided, but photographs of typical spectra are not reproduced. Comparatively little space is devoted to electronic band spectra, even a brief mention of which is considered by the author to be beyond scope of the book.

The title seems to be rather misleading: "outline" would surely be more apt than "elementary introduction." The book provides a useful short cut to present-day developments of the subject; the inclusion of microwave spectra will be especially welcome.

W. JEVONS

Reports on progress in physics. Volume XVI. Edited by A. C. STICKLAND. (London: The Physical Society, 1953.) Pp. 407. Price 50*s*.

The reviewer must take the full blame for the regrettable fact that his notice of the publication of Vol. XVI (1953) of *Reports on progress in physics* does not appear until after the

1954 volume has been issued! In any case this volume is a work on which no one nowadays can speak as an authority on all of the nine topics with which it deals. They cover neutron diffraction, physical properties and atomic arrangements in crystals, Raman effect in solids, paramagnetic resonance, semiconductors, circuit elements, electrical discharges, fluctuation theory in physical measurement, cosmology and the new unstable cosmic-ray particles.

If a member of the Institute, say one in the applied field, wishes at the end of a busy day to find out what is happening in realms of physics outside of his normal activities, he will find here authoritative articles, critical and fully documented. But he must not expect them to be easy reading. They do not and are not intended to give the lighter touch of the short more popular summaries that appear regularly elsewhere.

Some of the articles in this volume refer particularly to recent reviews of progress in a given field and then set out to survey subsequent developments under particular selected headings. An excellent one by Llewellyn Jones on electrical discharges is an example. Similarly in the article on semiconductor circuit elements by Blakemore, De Barr and Gunn discussion is limited mainly to silicon and germanium, reviewed in the light of recent advances in the theory of conduction mechanism.

Thus each article is a work of reference useful to anyone wishing for whatever purpose to obtain detailed knowledge of recent progress in a given direction; and each year we are grateful to the authors and the Physical Society for supplying them.

A. M. TYNDALL

Flow properties of disperse systems. Edited by J. J. HERMANS. (Amsterdam: North-Holland Publishing Co., 1953.) Pp. xi + 445. Price 70*s*.

This further volume in the series on the rheology of natural and synthetic substances attempts to treat the deformation and flow of disperse systems—suspensions, emulsions, gels, protein and polymer solutions, liquid sprays, atomized liquids, foams, smoke and powders—in a connected sequence. Since the behaviour of a disperse system depends on what is done with it, the expert authors of the ten chapters were faced with a considerable task, not made any easier by the confused terminology of the original papers. In such a wide subject, much has necessarily been omitted, but the scientist who has to examine disperse systems will find the book invaluable, not only as a source of references, but also in the analogies that are suggested by the juxtaposition of such diverse items, as, for example, suspensions of large particles, and polymer solutions.

The first three chapters on suspensions, emulsions and gels give a commendably clear account of recent experimental evidence and its interpretation. There follows a mainly theoretical account of suspensions of macromolecules (regarded as rigid bodies) where Brownian motion and the laws of hydrodynamics are both operative. The application to the determination of the size of macromolecules by means of viscosity measurements is discussed critically. This discussion is completed in the next chapter which provides an analysis of dilute solutions of flexible chain molecules.

In two chapters the formation, the droplet size distribution and evaporation of coarse and fine liquid sprays are discussed. As in the remainder of the book, the theory and experimental results are carefully exhibited. There follows a short account of foams.

The last two chapters are somewhat different. That on smokes—of which the essential feature is the instability of a gaseous dispersion—covers their formation, coagulation, flow and size determination. Throughout, the "aging" of a smoke by sedimentation, diffusion, coagulation and evaporation is a central theme. That on powders deals with the packing and flow of granular masses, for which the dispersing medium is only of secondary importance.

The book is clear, both in what is said and how it is printed. If there is any defect, it lies in the attempt to put a quart into a pint pot, but this is more than offset by the graphic picture of disperse systems that emerges from its pages.

R. L. BROWN

Low temperature physics. By CHARLES F. SQUIRE. (London: McGraw-Hill Publishing Co. Ltd., 1953.) Pp. x + 244. Price 46*s*. 6*d*.

There has been a spectacularly rapid expansion of low temperature work in the United States since the end of the war. While barely half a dozen cryogenic laboratories existed in the U.S.A. during the 1930s, today their number is reckoned in scores. This trend can be wholly attributed to the development of the Collins helium liquefier, a commercially produced apparatus which enables any non-specialized laboratory, provided it has sufficient funds, to reach temperatures down to 1° K without even having to worry about liquid air supplies. Thus began a new cryogenic era in the U.S.A. and American low temperature physicists are justified when they refer to the pre-war years of low temperature physics as B.C. (before Collins).

This mushroom growth of cryogenic laboratories made the appearance of an up-to-date textbook on low temperature physics particularly welcome. It is fitting that the author of the first comprehensive book on this subject to be published in the United States should be Professor Squire, who made valuable contributions to many branches of low temperature physics and who built up and directs the small but flourishing cryogenic laboratory at the Rice Institute, Texas.

His book is based on a course of university lectures and this accounts both for its virtues and its shortcomings. One appreciates the fresh spontaneity characteristic of the spoken word but in a book that sets out to cover in 200-odd pages the wide field of low temperature physics one would prefer carefully composed exposition and judiciously balanced material. Several examples of lack of balance may be given. The third law of thermodynamics is hardly mentioned and nothing is said about its fundamental bearing on low temperature physics; yet nearly 30 pages are devoted to theories of the equation of state of gases. Or, to turn to a point of detail, surely a complete textbook derivation of $Tp^{(1-\gamma)/\gamma} =$ const. and of $C_p(\partial T/\partial p)_H = T(\partial V/\partial T)_p - V$ could have been omitted and the space thus saved devoted to a discussion of the actual efficiencies of the Claude and Linde gas liquefaction processes. And, finally, while a few remarks on air conditioning might be justified, as a brief diversion, in a lecture, a whole paragraph on this subject is unnecessary and out of place in such a book.

Professor Squire's volume has, however, the merit of being comprehensive and up-to-date. There are extensive chapters on both the experimental and theoretical aspects of liquid helium and of superconductivity and all topics of importance to low temperature physics are dealt with in the chapters on magnetism and on the solid state. Undoubtedly low temperature physicists will find much of interest in these pages and will regard it as a useful reference book. But it is doubtful whether one could recommend this book wholeheartedly as an introduction to low temperature research for the inexperienced and uncritical beginner.

N. KURTI

Notes and comments

Scientia Electrica

We have received the first three issues of a new Swiss periodical *Scientia Electrica*. The main purpose of this periodical (written in German) is the publication of some of the lectures given in the colloquia on "Modern problems of theoretical and applied electrical technology," held fortnightly at the Eidgenoessische Technische Hochscule in Zürich under the chairmanship of Professor M. Strutt, although it is intended to publish other papers as well. Publication (by S. Hirzel Verlag, Zürich) will be at irregular intervals, the first three issues having appeared over the period October 1953 to June 1954, the price per copy varying between *Fr.* 4 and *Fr.* 4.80. The Editor-in-Chief is Professor M. Strutt, and the Managing Editor is Dipl. Ing. E. Rohner, Gloriastrasse 35, Zürich 7. The scope of the new periodical may be seen from the following selection of titles from the first three issues: Transistors in audio amplifiers, by M. J. O. Strutt; Material transfer in electrical switch contacts, by G. J. Ekkers; Corona losses on high voltage lines, by O. E. Gerber; Absolute voltage measurements, by G. Induni.

The mechanism of phase transformations in metals

The Institute of Metals will hold a one-day symposium on the mechanism of phase transformations in metals, in London about November 1955.

The symposium will be in two parts, the morning session being devoted to transformations which are largely governed by thermal activation (nucleation and growth reactions) and the afternoon session to transformations in which thermal activation is less important (martensitic reactions).

Review papers are being invited to introduce each session, but in addition offers of short original contributions (up to 3000 words) from workers in the field of phase transformations will be welcomed. In these contributions emphasis should be laid on the experimental results and their interpretation rather than on details of experimental technique.

All persons wishing to submit a short communication of this kind are invited to send the title and a brief summary, as soon as possible, to the Editor, The Institute of Metals, 4 Grosvenor Gardens, London, S.W.1. Final scripts must be in the Editor's hands by 31 January, 1955.

The properties of tin

A completely revised edition of *The properties of tin* has been published by the Tin Research Institute. It aims at presenting all the published atomic and nuclear, physical, physico-chemical and metallurgical data concerning pure tin, but does not give data for tin alloys, which are covered in a separate publication of the Institute. The main divisions of the data are as follows: atomic and nuclear properties; line spectra; X-ray data; crystallography; optical properties; electrical and magnetic properties; thermal properties; mechanical properties.

The handbook is intended for libraries and research laboratories to whom it will be distributed, free of charge; for private readers there is a charge of *2s. 6d.* which defrays part of the printing costs. Copies of the handbook, which contains 55 pages and a useful index, may be obtained from Tin Research Institute, Fraser Road, Perivale, Greenford, Middlesex.

Film structure and adhesion

The 1955 biennial conference of the Oil and Colour Chemists' Association has been given the title *Film structure and adhesion* and will be held from Tuesday, 7 June, to Saturday, 11 June, 1955.

Papers for presentation are invited and communications should be addressed to Mr. P. J. Gay, Hangers Paints, Ltd., Stoneferry Works, Hull. Further information on the conference may be obtained from the General Secretary, Oil and Colour Chemists' Association, Memorial Hall, Farringdon Street, London, E.C.4.

Hydraulic transport of solids—a correction

On page 304 of our August issue it was stated that copies of the *Proceedings of a Colloquium on the Hydraulic Transport of Coal* could be obtained gratis from the National Coal Board, Scientific Department, Hobart House, Grosvenor Place, London, S.W.1. This is incorrect; the copies are 10*s.* 6*d.* each.

Journal of Scientific Instruments

Contents of the September issue

SPECIAL ARTICLE

Photometry of the microscope with special reference to projection. By F. E. J. Ockenden.

ORIGINAL CONTRIBUTIONS

Papers

A sensitive recording extensometer. By G. R. Hyde.
A simple infra-red grating spectrometer for use in analysis. By J. Gaunt.
A reflectometer for the assessment of surface texture. By J. Halling.
A Geiger counter X-ray crystal spectrometer. By P. J. A. McKeown and A. R. Ubbelohde.
The application and limitations of the edge-diffraction test for astigmatism in the electron microscope. By M. E. Haine and T. Mulvey.
The preparation of ultra-fine grain photographic emulsions. By B. H. Crawford.
A sensitive floor for examining the movement of traffic. By R. G. Bateson.
The preparation and properties of pressed alkali halide disks with special reference to their use in spectroscopy. By M. A. Ford and G. R. Wilkinson.

Laboratory and workshop notes

An attachment to a dial-gauge thickness meter for measuring average thickness. By R. F. Blackwell.
A variable resistance electrometer head. By H. House and M. J. Morant.
Ball valves. By B. B. Bach, J. V. Dawson and L. W. L. Smith.
A self-balancing thermistor bridge. By A. F. Standing.
A rotary scrubber for cleaning large quantities of mercury. By M. Gibson, T. Charnley and S. J. Lord.
The use of electrically conducting glass in cloud-chamber construction. By C. Bowness and N. Cusack.
A universal slit-drive mechanism for a single-beam infra-red spectrometer. By D. Chapman and E. T. Sanders.
A multiway valve for pressure connexions. By D. Firth.

NOTES AND NEWS

New books
New instruments, materials and tools
Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone*: Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price *2s. 6d.* including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone*: Monarch 7644.)

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

Summarized proceedings of a conference on X-ray analysis —Cambridge, April 1954

The 1954 Spring Conference of the X-ray Analysis Group of The Institute of Physics was held at the Cavendish Laboratory, Cambridge, on 1–2 April, and was concerned with new developments in crystal-structure analysis and their applications.

The first morning session of the conference consisted of papers describing recent experimental developments and the interpretation of the results obtained from them.

Dr. G. E. Bacon (Atomic Energy Research Establishment, Harwell) spoke on "Single-crystal analysis by neutron diffraction." Early neutron diffraction work was done with powdered materials but now that the very important influence of secondary extinction was well understood, single-crystal specimens are being widely used. Dr. Bacon mentioned some of the added difficulties imposed by the relatively low neutron flux available at the present time. A new neutron diffractometer is being constructed at Harwell which will give greater neutron flux at the specimen, and will permit a more accurate allowance for secondary extinction, since neutron wavelengths as low as 0·4 Å will be used. For the interpretation of the data, two dimensional Fourier series are almost invariably used. Such a series represents a projection of the scattering density of point nuclei under the influence of thermal vibration, and as viewed through a microscope of definite resolving power. The fact that the scattering cross-section of nuclei does not decrease with angle of diffraction, apart from the effect of thermal vibration, means that for the same accuracy of experimental measurement, neutron diffraction experiments can lead to a more accurate determination of atomic co-ordinates than would be possible with X-rays. Series-termination effects assume greater importance but can be corrected by the same methods as are used to correct electron density maps. Dr. Bacon illustrated these points by describing some beautiful work on the distribution of nuclei in potassium dihydrogen phosphate,[1,2] both at room and at liquid air temperatures. Not only had the hydrogen nuclei been located but the oxygen positions were determined with greater accuracy than was given by previous X-ray studies of this important ferroelectric material. In reply to questions, Dr. Bacon made it clear that the resolving power of the room temperature measurements was not great enough for it to be possible to decide between two possible idealizations of the distribution of protons in the hydrogen bond at room temperature, although the hydrogen atom was clearly not located midway between two oxygens at low temperatures.

Dr. J. A. S. Smith (University of Leeds) next spoke on "Nuclear magnetic resonance as an aid to structure analysis," and gave a brief outline of the principles and measurements involved, with particular emphasis on the resonance spectra of crystals. The application of the method which is possibly of greatest interest to crystallographers is to the measurement of the distance between hydrogen nuclei. In favourable circumstances this distance can be found perhaps twice as accurately as by the best diffraction studies, although it is a disadvantage that the actual positions of the hydrogen nuclei in the unit cell are not determined. It is especially interesting that the method has shown the presence of H_3O^+ in nitric acid monohydrate.[3] Dr. Smith also outlined results obtained by various research workers in investigations of ammonium chloride,[4] potassium hydrogen fluoride,[5] hydrazinium fluoride,[6] gypsum,[7] oxalic acid dihydrate[8] and lithium sulphate.[9] Professor Lonsdale (University College, London) asked whether the method could be used to decide whether the motion of the proton in a hydrogen bond, such as occurs in potassium dihydrogen phosphate, should be described as oscillation about one position of minimum potential, or transfer to such positions. Dr. Smith replied that if the frequency of transfer was in a suitable range, it could be detected.

In the final paper of the morning session, Dr. H. Lipson (College of Technology, Manchester) described some new developments in "Optical methods." The various steps used in determining a crystal structure by such methods were described. These are:

(1) The preparation of a weighted reciprocal lattice, in two or in three dimensions, from the measured X-ray intensities, and its interpretation in terms of known features of the crystal structure, such as long chains of atoms, or benzene rings. This gives the orientation of these prominent features of a molecule.[10,11]

(2) When the molecular orientation has been found, a new graphical method devised by Dr. C. A. Taylor is used to find the position or positions of these parts of the molecule. The positions chosen are those which account best for the zeros of the weighted reciprocal lattice. This method had been used to show that the structure of triphenylene found by Klug[12] was almost certainly incorrect.

(3) The correctness of the structure found by these two steps is then tested by comparing the transform of the projection of one unit cell, derived optically, with the relevant section of the weighted reciprocal lattice.

(4) Finally, the signs of the structure factors can be found by deriving optically the transform of the projected structure, and comparing it with that of the same structure, plus an "atom" at the origin. By the addition of the latter, positive regions of the transform are increased in area and in intensity, while negative regions are diminished. The structure can then be refined by conventional methods.

In opening the discussion, Professor Sir Lawrence Bragg (Davy–Faraday Laboratory, Royal Institution, London) spoke of the great improvements that had been made to this method in the course of the development from the "fly's eye" type of apparatus, but stressed the need for more quantitative measurements of the optically-derived transforms. A number of speakers stressed the need to be on the look out for false pseudo-homometric structures when using this, or indeed any, method of structure determination.

The afternoon session was concerned with developments in the theory of structure analysis, and the revolution that is being brought about in the practice of accurate structure analysis by the use of high speed digital calculating machines. Dr. W. Cochran (University of Cambridge) outlined the developments that have been taking place in "Direct methods," and stressed the practical value of the inequalities between structure factors, discovered by Harker and Kasper.[13] Application of these inequalities always gives the relation between the signs of three unitary structure factors U(h), U(h') and U($h + h'$) as

$$s(h)s(h') = s(h + h').$$

The explanation of the fact that this result is very frequently correct when inequalities do not require it, is probably to be found in Sayre's work[14] rather than in that of Harker and

Kasper. The use of sign relations to determine crystal structures has been "mechanized," and sets of signs for structure factors which are consistent with these relations to some specified extent can be found by the EDSAC (electronic delayed storage automatic calculator) in a very short time. An example of a comparatively simple problem which was solved in this way was briefly described. An extension of the method which makes use of the fact that certain structure factors are zero was mentioned, and one successful application described. Recent work on the theory of sign relations, by Hauptman and Karle[15] and by Woolfson[16] was briefly reviewed, but the claim by the former authors to have found a general solution of the phase problem for centrosymmetrical crystal structures was dismissed as unfounded.

Dr. A. S. Douglas (University of Cambridge) described how the EDSAC has been used to determine crystal structures from sign relations between the structure factors.[17] The equation $s(h)s(h') = s(h + h')$ states one of the following:—

$+ \times + = +$, $- \times - = +$, $+ \times - = -$, $- \times + = -$.

If we replace $+$ by zero and $-$ by unity, these equations are completely analogous to

$0 + 0 \equiv 0, 1 + 1 \equiv 0 \pmod 2, 1 + 0 \equiv 1, 0 + 1 \equiv 1.$

Thus a set of rather complicated equations can be replaced by a set of linear equations, and since they involve only the numbers 0 and 1, they are ideally suited for handling by a machine which works on the binary system. Further details of the process of finding on the EDSAC which set or sets of signs for a selected group of structure factors cause the sign relationship to hold within specified limits were given, and the results obtained on two test structures were shown to be satisfactory.

The next two papers described the use of the Manchester automatic computer for accurate structure analysis, that is, for the calculation of structure factors and various types of Fourier synthesis. Dr. F. Fowweather (College of Technology, Manchester) had as his title "Methods of calculating structure factors and Fourier syntheses" and he described a general-purpose programme for Fourier synthesis which can be applied to space-groups of any symmetry, including non-centrosymmetrical ones, without modification of routines. This obviates the need to check programmes for different space-groups. A synthesis which might have taken twenty hours to calculate by hand was completed in under one hour, including preparation and checking time. A general programme for structure factor calculations was also described. It has the disadvantage that atomic co-ordinates must be rounded off to the nearest 1/256. It is, however, more accurate than the method of Beevers and Lipson[18] on which it is based. The time required was five seconds per structure factor in a typical example.

Dr. P. J. Wheatley (University of Leeds) spoke on "The use of the Manchester machine for accurate structure analysis," and opened by saying that he did not intend to talk about what he had done for the machine, as the programmes had been prepared by Ahmed and Cruickshank,[19] but about what the machine had done for him. He described the use that has been made of the machine for three-dimensional refinement of approximate atomic co-ordinates, and possibly aroused the envy of those with no access to electronic computers by pointing out that tedious work requiring some months of hand computation could now be completed in as many days. The process involves computation of structure factors for all planes for which a value of the structure factor has been recorded, the use of these calculated structure factors and the observed structure factors to compute differential syntheses, and the combination of the results from these syntheses to obtain improved atomic co-ordinates, corrected for series-termination errors. The reliability index and the figures required to estimate the accuracy of the co-ordinates can be obtained at each stage of the refinement. Programmes for this iterative procedure have been prepared for five different space-groups and eighteen structures refined, using three-dimensional data, in about eighteen months. In the refinement of the structure of benzene, one complete cycle took only thirty-one minutes.

Dr. A. D. Booth (Birkbeck College, London) opened the discussion and gave some times for structure factor calculations on the Birkbeck machine. He described the practice of repeating a calculation for checking purposes as of little value.

Dr. D. W. J. Cruickshank (University of Leeds) gave the final paper of the day's session on "Anisotropic thermal motions and molecular charge distributions" and began by pointing out that the investigation of such factors was not merely a matter of computation, but depended primarily on the development of accurate methods of intensity measurement. In certain cases, however, the effect of non-uniform thermal vibration on the electron distribution is so large as to be easily detected, for example in geranylamine hydrochloride. A reinvestigation, using the original experimental data,[20] has shown that values of the temperature factor B in this case vary between $4{\cdot}0$ Å^2 for an atom at one end of the molecule to $11{\cdot}4$ Å^2 for one at the other end. When this effect is allowed for, the measured bond lengths show no abnormal features. In other cases, for example naphthalene, the thermal motion can be described as an isotropic oscillation of each atom, plus an oscillation of the molecule as a whole about a fixed centre. This results in each atom having an anisotropic temperature factor, that for atoms on the outside of the molecule being largest. Before the charge distribution in a molecule can be interpreted in terms of bonding electrons etc., such effects must be allowed for. A programme which will refine atomic temperature factors is being prepared for the Manchester computer. The criterion adopted is the equality of calculated and observed electron densities, as well as their second derivatives, at the centre of each atom. Dr. Cruickshank next discussed the use of maps of $(\rho_0 - \rho_c)$ to investigate the difference between the charge distribution in a molecule and that in a corresponding superposition of isolated atoms. The redistribution of charge on bonding may be difficult to distinguish from the effects of thermal vibration. An approximate calculation can be made which makes no allowance for this factor.

If the electron density in two non-bonded atoms is

$$\rho_c = \psi_a^2 + \psi_b^2,$$

on bonding it becomes

$$\rho_0 = \frac{\psi_a^2 + 2\psi_a\psi_b + \psi_b^2}{1 + S}$$

where $S = \int \psi_a \psi_b d\tau$, the overlap integral. Calculations have been made for the σ-bonds in graphite on this basis and show values of $(\rho_0 - \rho_c)$ in the plane of a layer of graphite which vary from $+0{\cdot}215$ e Å^{-3} at the centre of a bond to $-10{\cdot}6$ e Å^{-3} at an atomic centre. Even in a parallel plane at some distance the σ-density is found to outweigh the π-density, except immediately above the carbon atoms. Results obtained by March[21] agree with this conclusion. In a typical case ($B = 5{\cdot}5$ Å^2) the temperature factor reduces the difference density at an atomic centre to $-0{\cdot}45$ e Å^{-3} without appreciably affecting that in the bond. The difference density will be very difficult to measure at an atomic centre, however, as it is so greatly dependent on the value of B used,

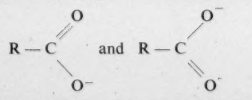
but the prospects for the measurement of the bond density by X-ray methods are fairly good.

The evening discourse by DR. R. C. EVANS (University of Cambridge) took the form of "Editorial reflexions" and was based on the experience of six years as a co-editor of *Acta Crystallographica*. Dr. Evans observed that publication was the one subject in which all scientists were interested; so much so that it was estimated that scientific papers were now being published at the rate of one every six seconds. This might be held to show the need for more journals devoted to a limited field of investigation and the comparative uselessness of journals which accepted articles on any and every branch of science. Passing from the general to the particular, the speaker then outlined some of the guiding principles of an editor—his responsibility to the subject, to the authors and to the readers. While authors were active in defence of their interests, readers were less so, and Dr. Evans asked them to take the opportunity of expressing their views. On the whole, papers submitted to *Acta Crystallographica* were not written in a deplorable literary style. Papers from American authors were almost invariably good as far as technical preparation was concerned, those from German authors in particular were often very bad in this respect. Some amusing examples of language difficulties were instanced. "Diffraction by little Bragg angels" and "Please thank the referees for their suggestions. I would like to execute them," are worthy of a wider public. The speaker gave some advice on the preparation of a paper for publication, and finally some details of the economics of scientific journals in general, and *Acta Crystallographica* in particular.

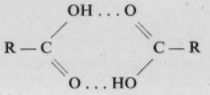
In a short discussion which followed this lecture, Professor Bragg suggested that it should be published, possibly after a referee had suggested how it might be reduced in length! Various speakers put forward suggestions for reducing the cost of printing numerical tables, and some suggested that there was no need to publish tables of observed and calculated structure factors in papers describing a crystal structure analysis. On this latter point it was clear that there was a wide divergence of views.

The papers given on the second day of the conference mostly dealt with results which have been obtained in various accurate structure analyses. PROFESSOR J. M. ROBERTSON (University of Glasgow) outlined "Some precise determinations of carboxylic acid structures." Recently the structure of ethynylacetic acid $HC{\equiv}C.CH_2.COOH$ has been determined, chiefly in order to find, if possible, the electron distribution in acetylenic hydrogen, but the results have also provided interesting data on the dimensions of the carboxyl group. Additional information has also come from a three-dimensional refinement of succinic acid, $COOH(CH_2)_2COOH$, using the Manchester computer.

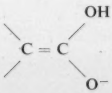
An ionized carboxyl group may be expected to be completely symmetrical, as the two resonating structures



are completely equivalent. This agrees with results obtained in investigations of DL-serine, L-threonine and hydroxy-L-proline. It is, however, rather puzzling that certain other structures such as DL-alanine, β-glycylclycine and glycyl-L-asparagine, which are presumably still ionic from their chemical constitution, exhibit a significant difference between the two C–O distances. It has been suggested by Donohue[22] that one resonance structure may be stabilized by multiple hydrogen bonding of one of the oxygens, but the speaker doubted whether this was an adequate explanation. Molecules containing a carboxyl group may be expected to form hydrogen bonds such as in



and this may have some influence on the dimensions of the carboxyl group. There is, however, no convenient way in which the carboxyl group can be studied in isolation. In β-succinic acid, salicylic acid and benzoic acid the dimensions of the carboxyl group are very similar but differ from those found in oxalic acid dihydrate, α-oxalic acid and α-ethynylacetic acid, particularly in that in the former class of compounds the C–C bond is about 1·47 Å long, whereas in the latter it is close to the single-bond value of 1·54 Å. At least in the case of the aromatic compounds, the contribution of the structure



provides an adequate explanation of this fact. Professor Robertson stressed that the establishment of these, and similar points, depended on the developments in computational techniques and in the theory of accurate structure analysis which had taken place in the last few years. In the discussion, PROFESSOR COX (University of Leeds) suggested that when a compound contained a large number of carboxyl groups, at least one might be found which made no hydrogen bonds. The influence of these bonds on the carboxyl group could then be deduced. DR. DUKE (Explosives Research and Development Establishment, Waltham Abbey, Essex) gave a short account of an analysis of oxamide. Very high accuracy had been achieved in this case by using molybdenum characteristic radiation, and because the atoms have a very small temperature factor.

DR. H. P. STADLER (King's College, Newcastle) illustrated "The application of Fourier transforms to structure analysis" by two examples. In flavanthrone, the projection on the *b*-face of the unit cell contains all molecules in the same orientation, so that an immediate comparison of the weighted reciprocal lattice with the transform is possible, but in pyranthrone two transforms have to be combined to allow for the two different orientations of the molecules in projection. In the latter case, there are a number of pseudo-homometric structures which, although incorrect, and occasionally obviously so from a consideration of minimum intermolecular distances, are, nevertheless, not immediately ruled out by the observed X-ray intensities. Dr. Stadler also discussed the calculation of a transform by Beevers-Lipson strips with special reference to errors introduced by neglecting differences in bond lengths and atomic numbers. In the discussion DR. KLUG (Birkbeck College, London) and DR. WOOLFSON (University of Cambridge) both mentioned other instances where pseudo-homometric structures had been found, and where it had proved difficult or impossible to find the correct solution.

DR. P. J. BLACK (University of Cambridge) gave the last paper of the session on "Electron distribution in intermetallic compounds." Work on the correlation of phase diagrams

and on the Brillouin zones of aluminium-rich compounds has led to the suggestion that in such compounds, electrons are absorbed to fill vacancies in the 3*d* shells of the transition metals.[23] Attempts have been made to observe the effect directly on maps showing $(\rho_0 - \rho_i)$. Dr. Black examined critically the supporting evidence obtained by some of these studies[24, 25, 26] and showed that errors in the Viervoll and Ogrim scattering factors[27] largely invalidate the X-ray evidence for the supposed electron transfer. Some conclusions by Lonsdale and Bijvoet[28] about the possibility of detecting electron transfer by X-ray methods were shown by the speaker to be over-pessimistic in the general case.

The last two papers given to the conference took the form of general surveys. PROFESSOR E. H. WIEBENGA (Rijks University, Groningen, Holland) spoke on "The application of modern structure analysis to organic chemistry." The first applications of X-ray analysis were mostly to inorganic chemistry, although by now the crystal structures of some 350 organic compounds are known. In the past, structure analysis has been in fact less necessary (as well as more difficult), for organic than for inorganic compounds, but a change is being rapidly brought about by the development of theoretical organic chemistry.

X-ray methods have been used to distinguish between isomers in a number of instances. For example, to determine the structures of tropine and of isoleucine the "heavy atom" method has been used. Two possible tautomeric forms of α-pyridone have been distinguished by locating the hydrogen atoms in this compound,[29] while the absolute configuration of d-rubidium tartrate has been settled by making use of the fact that Friedel's law is not obeyed in certain circumstances.[30] Sometimes the interest has been in the "architecture" of the crystal, for example in the case of clathrates and other intermolecular compounds, but most present-day work is concerned with the accurate measurement of bond lengths and bond angles. For these results to be usefully compared with theoretical calculations, accurate measurement of the X-ray intensities and allowance for systematic errors such as those associated with Fourier-series termination, is essential. Further improvements will no doubt be brought about by measuring all the X-ray intensities and not merely those from one zone by means of counter diffractometers, and by making such measurements at low temperatures. The structure analysis of complex organic compounds will probably continue to depend on the preparation of suitable derivatives containing a heavy atom. Many of the examples which Professor Wiebenga used to illustrate his points were selected from work which is being carried on in Holland at the present time.

DR. A. F. WELLS (Imperial Chemical Industries Ltd., Manchester) followed with a paper on "The application of modern structure analysis to inorganic chemistry." For many years little attention has been paid to the geometrical aspects of the crystal chemistry of inorganic compounds. There are many interesting relationships between the structures of different compounds which do not appear to have been noted. These relationships emerge from a study of systems of points connected up to form networks which can either be finite, or can extend indefinitely in one, two or three dimensions. For example, the structure of one form of P_2O_5 can be idealized to that of a system of points, each connected to three other points. From this point of view, the structure of P_2O_3 is similar. Hydrogen peroxide may be regarded as an example of the analogous three-connected system in three dimensions. The three polymorphs of P_2O_5 provide examples of three different types of three-connected systems, while certain silicates provide examples of four-connected systems. The structures proposed for the inert gas hydrates are very plausible when examined from this point of view.

Again, the removal of $\frac{1}{2}$Na from certain layers of the NaCl structure produces the structure of $CdCl_2$, while their removal in a different way leads to the structure of $Cu_2(OH)_3Cl$. These geometrical relations are completely hidden when the structures are described in the conventional manner, in terms of space group symmetry, or structure type.

In the discussion on these papers, PROFESSOR C. MACGILLAVRY (University of Amsterdam, Holland) suggested that geometrical considerations similar to those discussed by Dr. Wells might explain why some compounds, particularly among the sugars, are so difficult to crystallize. Perhaps in some instances a hydrogen-bond network could not be formed. A number of speakers commented on the value of making intensity measurements with molybdenum characteristic radiation at low temperatures, for accurate structure analysis.

W. COCHRAN

REFERENCES

(1) BACON, G. E., and PEASE, R. S. *Proc. Roy. Soc. A*, **220**, p. 397 (1953).
(2) BACON, G. E., and PEASE, R. S. *Nature* [*London*], **173**, p. 443.
(3) RICHARDS, R. E., and SMITH, J. A. S. *Trans Faraday Soc.*, **47**, p. 1261; **48**, p. 675 (1951).
(4) GUTOWSKY, H. S., PAKE, G. E., and BERSOHN, R. *J. Chem. Phys.*, **22**, p. 643 and p. 651 (1954).
(5) WAUGH, J. S., HUMPHREY, F. B., and YOST, D. M. *J. Phys. Chem.*, **57**, p. 468 (1953).
(6) RICHARDS, R. E. Private communication.
(7) PAKE, G. E. *J. Chem. Phys.*, **16**, p. 327 (1948).
(8) ITOH, J., KUSAKA, R., KIRIYAMA, R., and YABUMOTO, S. *J. Chem. Phys.*, **21**, p. 1895 (1953).
(9) SOUTIF, M., DREYFUS, B., and AYANT, Y. *C.R. Acad. Sci.* [*Paris*], **233**, p. 395 (1951).
(10) HANSON, A. W., LIPSON, H., and TAYLOR, C. A. *Proc. Roy. Soc. A*, **218**, p. 371 (1953).
(11) KENYON, P. A., and TAYLOR, C. A. *Acta Cryst.*, **6**, p. 745 (1953).
(12) KLUG, A. *Acta Cryst.*, **3**, p. 165 (1950).
(13) HARKER, D., and KASPER, J. S. *Acta Cryst.*, **1**, p. 70 (1948).
(14) SAYRE, D. *Acta Cryst.* **5**, p. 60 (1952).
(15) HAUPTMAN, H., and KARLE, J. *A.C.A. Monograph No.* 3 (Wilmington, Delaware: The Letter Press, 1953).
(16) WOOLFSON, M. M. *Acta Cryst.*, **7**, p. 61 (1954).
(17) COCHRAN, W., and DOUGLAS, A. S. *Nature* [*London*], **171**, p. 1112 (1953).
(18) BEEVERS, C. A., and LIPSON, H. *Acta Cryst.*, **5**, p. 673 (1952).
(19) AHMED, F. R., and CRUICKSHANK, D. W. J. *Acta Cryst.*, **6**, p. 765 (1953).
(20) JEFFREY, G. A. *Proc. Roy. Soc. A*, **183**, p. 388 (1945).
(21) MARCH, N. H. *Acta Cryst.*, **5**, p. 187 (1952).
(22) DONOHUE, J. *J. Phys. Chem.*, **56**, p. 502 (1952).
(23) RAYNOR, G. V. *Progr. Metal Phys.*, **1**, p. 1 (1949).
(24) DOUGLAS, A. M. B. *Acta Cryst.*, **3**, p. 19 (1950).
(25) ROBINSON, K. *Acta Cryst.*, **5**, p. 397 (1952).
(26) NICOL, A. D. I. *Acta Cryst.*, **6**, p. 285 (1953).
(27) VIERVOLL, H., and OGRIM, O. *Acta Cryst.*, **2**, p. 277 (1949).
(28) BIJVOET, J. M., and LONSDALE, K. *Phil. Mag.*, **44**, p. 204 (1953).
(29) PENFOLD, B. R. *Acta Cryst.*, **6**, p. 591 (1953).
(30) BIJVOET, J. M., PEERDEMAN, A. F., and VAN BOMMEL, A. J. *Nature* [*London*], **168**, p. 271 (1951).

A method of examining selected areas of surfaces using replicas and the electron microscope*

By G. R. Booker, B.Sc., A.Inst.P.,† Research Department, Metropolitan-Vickers Electrical Co. Ltd., Manchester

[*Paper received* 15 *February*, 1954]

A simple method has been devised for examining selected areas of surfaces using wet-stripped replicas. The selected area of the surface is identified by fine scratches in the immediate vicinity. The corresponding portion of a replica is mounted over a 0·015 in. diameter hole in a copper disk and viewed in an electron microscope operating with a specimen screening aperture. The method has several advantages over methods hitherto used, e.g. it has extremely wide application as it does not depend on surface structure for identification, the positioning of the replica on the copper disk is rapid and can be correctly performed each time, and an uninterrupted area of 0·015 in. diameter can be observed.

After the examination of the surface of a solid specimen with an optical microscope, it is often desirable to examine particular features in greater detail in the electron microscope. This generally necessitates the use of a replica, which introduces two main difficulties:

(*a*) identification of the area on the replica, and
(*b*) favourable mounting so that the area can be imaged.

Several solutions to the problem have been proposed, each with limitations.

Hyam and Nutting[1] have described a method for dry-stripped replicas in which the grid is carefully placed on to the required area of the replica before stripping by using a micro-manipulator attached to an optical microscope. A somewhat similar method was used independently by Radavich.[2] The disadvantage of these two methods is that they usually necessitate thick replicas and produce strains in the replica.

Agar and Revell,[3] using wet-stripped replicas, mark a particular area by fine scratches and mount the replica in the normal manner. Several replicas are sometimes required before a favourable positioning for observation of the area occurs. This is slow and tedious and the surface of the specimen may deteriorate before success is attained.

Recently, Nankivell[4] has described an alternative method for wet-stripped replicas. The method relies on use of the structure of the replica, as observed under an optical microscope, to position a particular area over a square grid opening of about 0·003 in. side. This method has very limited application. For metals, a deep etch is necessary, often causing replica artifacts. Even with a deep etch many metal structures would be difficult to identify, e.g. bainite steel, whilst lightly-etched metal surfaces or smooth surfaces would be impossible to deal with.

A simple method has been devised which avoids these limitations. It uses the method of Agar and Revell for identification and substantially the method of Nankivell for mounting. There are, however, some important differences that justify a description of the method.

METHOD

Scratches about 2 μ wide are made on the specimen surface by the use of a Vickers micro-hardness tester. A typical pattern is reproduced in Fig. 2(*a*). The features of interest lie within the four enclosed areas, each of which can be identified by its shape and relation to the others without reference to structure.

A replica is made by the modified process for wet-stripped backed replicas described by Agar and Revell,[5] the replica having been scored so that the scratched area is roughly central. In its final form it is floated on to the surface of water held in a glass dish, *A* in Fig. 1. A three-legged copper

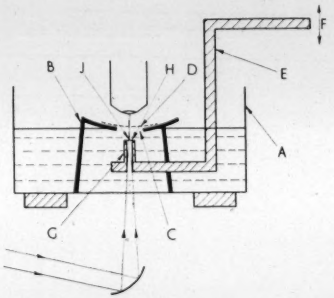


Fig. 1. Apparatus showing replica ready for mounting

stool *B*, with a curved seat containing a central hole *C* about 0·25 in. in diameter, stands completely immersed in the dish. The replica *D* is positioned over the hole *C* and water slowly removed from the dish using a micro-pipette, until the replica is located with its outer portion resting on the stool seat, and the central portion lying across the hole with its lower side in contact with the water.

When the replica is in this position it is firmly held by the stool and does not move relative to the stool as a result of adventitious disturbances. A stool with a flat seat was used initially but it was found that surface tension effects often made location difficult. With a curved seat, however, the water is removed very smoothly through the central hole of the curved surface and no such difficulty arises.

The dish is then placed on the stage of an optical microscope and a brass arm *E* lowered into the water by means of a fine vertical rack-and-pinion movement *F* attached to a heavy stand resting on the bench. The arm *E* holds a small brass tube *G* to which is attached, by a minute quantity of adhesive,

* This paper was presented at the meeting of the Electron Microscopy Group of The Institute of Physics held in London, November 1953.

† Now at RTSC Laboratories, Whitchurch, Aylesbury, Bucks.

an $\frac{1}{8}$ in. diameter, 0·001 in. thick, copper disk H with a central hole J of 0·015 in. diameter. The aperture disk H should be flat and approximately horizontal, i.e. parallel to the replica.

The replica is viewed in transmission at about ×250. Under these conditions the scratches are readily visible, allowing the scratched area to be positioned in the centre of the field of view by means of the stage controls.

The brass tube G is moved under the replica from the side

(*a*) (*b*)

Fig. 2. (*a*) Optical micrograph showing marked area of a metal surface. Magnification ×160. (*b*) Optical micrograph of a replica of the metal surface in position over the 0·015 in. diameter hole in an aperture disk. Magnification ×160

and raised by the movement F until the aperture disk H is close to the replica. The hole J can then be seen through the microscope and is placed centrally into the field of view, i.e. symmetrically surrounding the scratch pattern, by moving the stand on the bench.

The arm E is slowly raised until the aperture disk H touches the replica, any slight misalinement of the aperture and the scratched area being corrected during the process with the stage controls.

Further raising lifts the replica completely from the stool, causing the outer portion of the replica, surrounding the area actually in contact with the disk H, to wrap itself round the outside of the tube G. The raising process is continued until the stool, resting on the arm E, is lifted out of the water. Water adhering to the tube and arm is absorbed with filter paper, the stool removed, and the arm holding the aperture disk and replica placed in a desiccator. When completely dry the aperture disk can be safely detached from the tube G by a razor blade, and the specimen is then ready for examination. An optical micrograph of the replica in position on the aperture disk is shown in Fig. 2(*b*).

The electron microscope used is fitted with a specimen screening aperture of the form described by Page and Agar.[6] This enables relatively thin replicas (showing a blue interference colour) spanning a 0·015 in. diameter hole to be viewed at high intensity without damage. Although not so convenient or effective, it is sometimes possible, in the absence of a specimen screening aperture, to use a suitably reduced condenser aperture successfully.

CONCLUSIONS

1. The method can be applied to any surface that can be scratched with a diamond and from which replicas can be obtained using the wet-stripping process. It is independent of surface structure for identification purposes.

2. The surface can be treated in diverse ways between replicas, and the particular area can still be correctly positioned as long as the treatment does not destroy or seriously reduce the depth of the scratches.

3. The ability to view an uninterrupted area of 0·015 in. diameter is very useful for many investigations.

4. Wet-stripped replicas made by any process, and not necessarily by the process suggested in the paper, can be used.

5. Once the replica has been positioned on the aperture disk no further operations are performed, so that risk of subsequent movement is avoided.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. Willis Jackson, Director of Research and Education, and Mr. B. G. Churcher, Manager of Research Department, Metropolitan-Vickers Electrical Co. Ltd., for permission to publish this paper.

REFERENCES

(1) Hyam, E. D., and Nutting, J. *Brit. J. Appl. Phys.*, **3**, p. 173 (1952).

(2) Radavich, J. F. Private communication.

(3) Agar, A. W., and Revell, R. S. M. Paper read at the Electron Microscopy Conference, Bristol, September 1952.

(4) Nankivell, J. F. *Brit. J. Appl. Phys.*, **4**, p. 141 (1953).

(5) Agar, A. W., and Revell, R. S. M. *Brit. J. Appl. Phys.*, **2**, p. 8 (1951).

(6) Page, R. S., and Agar, A. W. *J. Sci. Instrum.*, **31**, p. 27 (1954).

A potentiostat for corrosion study

By M. H. Roberts, B.Sc., A.Inst.P.,* Brown-Firth Research Laboratories, Sheffield

[*Paper first received* 3 *March and in final form* 5 *April*, 1954]

The potential difference between a standard electrode and a test electrode in an electrolytic cell forms the input to a thermionic d.c. amplifier. The output current of the amplifier flows between the test electrode and a third electrode of platinum. Any change in potential of the test electrode alters the output current to reduce the charge, the system forming a closed-loop proportional controller. The two-stage amplifier is symmetrical, so that zero or reverse current can be obtained without loss of sensitivity. An output of 20 mA can be supplied with less than 100 mV variation of electrode potential.

In studying the relationship between electrode potential and corrosion rate, it is desirable to carry out experiments either at constant current (e.g. by applying a high voltage through a high resistance) or at constant potential. In some experiments it proves practicable to apply the desired constant potential difference, derived from a potential divider of low resistance, directly between the test electrode and a platinum electrode used as standard. However, in many instances the current flowing through the platinum electrode alters its potential, so that it is desirable to have a separate standard electrode (through which no current is passed) for the reference potential, while the electrolytic corrosion current flows between the test electrode and the platinum electrode. This arrangement calls for some device which will respond to any change in the potential difference between standard and test

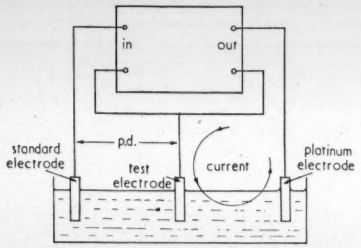


Fig. 1. Outline of system for studying the relationship between electrode potential and corrosion rate

electrodes, without consuming current from the former, and which will vary the current between platinum and test electrodes in such a way as to correct the original deviation. Fig. 1 outlines the system.

EARLIER WORK

An electronic circuit for achieving this result was first used by Hickling,[1] who named it a "potentiostat." The input potential was applied to the control grid of a thyratron, so that a rise in potential caused the thyratron to fire. The rectified output voltage of the thyratron was smoothed and applied as bias to the output valve, varying its anode current. The discontinuous response of the thyratron necessitated a "hunting" or saw-tooth oscillation of the system to maintain control. The output current was unidirectional, being the anode current of the output valve, and performance deteriorated as cut-off was approached. It was therefore decided to design a circuit in which the response was continuous and the output current reversible through zero without loss of sensitivity.

* Now at Joseph Lucas Ltd., Birmingham.

SPECIFICATION OF PERFORMANCE

The performance required is that the potential should be kept within 100 mV of the set value, and that an output current of up to 20 mA may be required, depending on the electrode area. Thus a d.c. amplifier, having an overall transconductance of 200 mA/V or more, with one input and one output terminal common, will provide satisfactory control, the closed loop feedback system being completed through the electrolytic cell. Since the amplifier functions as a simple proportional controller, there will be an "offset" error of 100 mV when the current demanded changes by 20 mA in the above system. As there may also be some gradual drift in the first stage, slightly greater sensitivity is preferable, to allow some margin for drift.

GENERAL DESIGN

This sensitivity can be obtained in two stages, the first using a pair of pentodes giving a gain of about ×25, while the output stage uses high-slope tetrodes as cathode followers. Fig. 2 shows the circuit. Both stages are in push-pull arrangement as this gives many advantages. Drift caused by varying

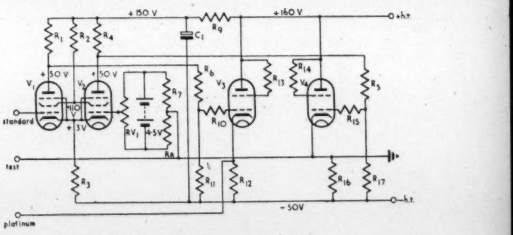


Fig. 2. Circuit of potentiostat

R_1 = 220 kΩ
R_2 = 220 kΩ
R_3 = 68 kΩ
R_4 = 220 kΩ
R_5 = 470 kΩ
R_6 = 470 kΩ
R_7 = 10 kΩ
R_8 = 10 kΩ
R_9 = 10 kΩ
R_{10} = 1 kΩ
R_{11} = 470 kΩ
R_{12} = 2 kΩ, 3 W
R_{13} = 100 Ω
R_{14} = 100 Ω
R_{15} = 1 kΩ
R_{16} = 2 kΩ, 3 W
R_{17} = 470 kΩ
RV_1 = 25 kΩ
C_1 = 8 μF
V_1 = Z729
V_2 = Z729
V_3 = KT61
V_4 = KT61

supply voltages is reduced; the current drawn from the h.t. supply remains practically constant; and the characteristic curve of output current *versus* input voltage is symmetrical about the zero current axis. Variable bias on the alternate input grid can be used for offsetting the working point to the desired potential.

The h.t. supply is not earthed except *via* the cathode of one output valve. The test electrode is connected to the common earth of the input and output circuits and the standard electrode to the live input terminal. As the resistance

associated with the type of standard electrode used is not very high, electrostatic screening of the electrode and the lead to it has not been found necessary, but probably would be on a system using a very high resistance (e.g. glass) electrode.

TESTING

The operation of the potentiostat should be checked by measuring the input-output characteristic, either by applying various steady input potentials and measuring the output current, or by applying an alternating input voltage and plotting the output across a small load resistance against it on a cathode-ray oscillograph by connecting X and Y to input and output. Balance between the two sides should be set by adjusting the bias on V_2, while the input to V_1 is shorted to earth.

MODIFICATIONS FOR GREATER OUTPUT CURRENT

The instrument described gives adequate output for research on small samples. While the order of output voltage needed is only about 1 V, the current required will be proportional to the electrode surface area, and the inefficiency of the hard valve, when working into such a low resistance load, becomes uneconomic if a new method of electrolytic protection for stainless steels, demonstrated on pilot plant scale at the Society of Chemical Industry's Corrosion Group Conversazione (January, 1954), is to be used on a large scale. Higher output currents can be provided by a high power servo-operated potential divider, a variable field d.c. generator or a high current phase-sensitive rectifier. The latter has been used by the author to provide 0·5 A output from a single valve, but the smoothing presents a problem. The input potential change can be converted to a.c. by a balanced modulator circuit, or by a vibrating contact converter, which eliminates drift. Some arrangements of this type have been described,[2, 3] but are more complex than is necessary for the small-scale research experiments for which the above instrument was designed.

REFERENCES

(1) HICKLING. *Trans Faraday Soc.*, **38**, p. 27 (1942).
(2) LINGANE. *Industr. Engng. Chem.* (Analyt. Ed.), **17**, p. 332 (1945).
(3) PALMER and VOGEL. *Analyst*, **78**, p. 429 (1953).

The glass sealing properties of titanium and zirconium

By H. RAWSON, M.Sc.Tech., and E. P. DENTON, B.Sc. Tech., Research Laboratory, The British Thomson-Houston Co. Ltd., Rugby

[*Paper received* 5 *January*, 1954]

Titanium and zirconium may be sealed to standard sealing glasses without the need for any special sealing techniques, to give strong vacuum-tight seals with low stresses in the glass. If a flame-sealing method is used it is necessary to keep the sealing time as brief as possible, since the expansion characteristics of both metals change appreciably on prolonged heating in air.

During the last few years titanium and zirconium have been made commercially available in the form of tubing and wire. It was known that the linear thermal expansion coefficients of these metals lay in the range 50–100 $\times 10^{-7}$ per °C and it was considered likely, therefore, that they could be sealed to standard sealing glasses.

THERMAL EXPANSION PROPERTIES

Fig. 1 shows the results of thermal expansion measurements made on rods of zirconium and titanium obtained from various sources. Fig. 1 also shows the effect of heating the rods for 5 min in air at 1100° C, a treatment chosen to simulate the most extreme conditions likely to be experienced in making a glass-to-metal seal. The curves clearly show that the expansion characteristics of zirconium and titanium vary quite markedly, depending on the source of the material and the heat treatment given to it, and this point should always be borne in mind when making seals using these metals.

GLASS SEALING PROPERTIES

Satisfactory vacuum tight seals were made on samples of zirconium and titanium wire using a number of glasses. The wire needed no special treatment prior to sealing other than rubbing with fine emery paper to remove the drawing marks and then heating for a few seconds in the flame. Fig. 2 shows a number of seals made on both wire and tubing.

In order to study the thermal expansion match in the various metal-glass combinations, the stresses in specially prepared bead seals were measured over the whole temperature range from the annealing temperature of the glass down to room temperature by a photoelastic technique.[1] Fig. 3 shows the type of specimen used. The retardation was measured at the glass-metal interface, the light beam passing axially

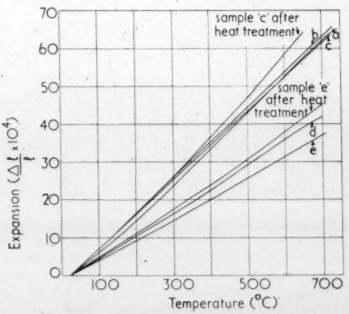


Fig. 1. Thermal expansion curves of various samples of titanium and zirconium

a, titanium from source 1; *b*, titanium from source 2; *c*, titanium from source 3; *d*, zirconium from source 4; *e*, zirconium from source 5.

through the specimen, so the retardation measured was proportional to the sum of the hoop stress, P_θ, and the radial stress, P_r. The retardation-temperature curves for the three best sealing combinations (C.40 to zirconium and C.77 and

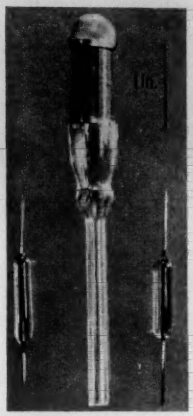


Fig. 2. Seals of C.40 glass to zirconium wire and tubing

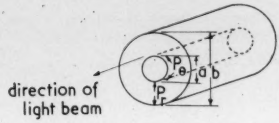


Fig. 3. Type of bead seal used in determining retardation-temperature curves

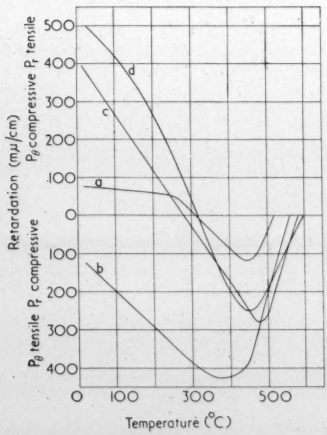


Fig. 4. Retardation-temperature curves for bead seals of:
a, C.40-zirconium (source 5) b/a = 7·6; *b*, C.77-titanium (source 3) b/a = 7·5; *c*, C.78-titanium (source 3) b/a = 7·9; *d*, C.9-tungsten, b/a = 6·2.

C.78 to titanium) are shown in Fig. 4. C.40 glass is a standard glass for sealing to iron–nickel–cobalt alloys, such as Nilo K, and C.77 and C.78 are standard glasses, developed by Mr. J. E. Stanworth of this laboratory, for the production of graded seals. In the same figure is plotted, for comparison purposes, a retardation-temperature curve for C.9-tungsten, a standard sealing combination.

THE MECHANICAL STRENGTH OF GLASS SEALS

Fig. 5 shows the type of test used to determine the adhesion between the glass and the metal. The test seal consisted of a

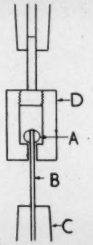


Fig. 5. Technique used for determining the strength of glass to metal seals

bead of glass, *A*, about 6 mm in diameter, melted down on to a 7·5 cm length of 1 mm diameter wire, *B*. The wire was held in the lower jaw *C* of a tensile testing machine and was pulled downwards, so that the chuck *D* tended to shear the glass bead off the wire at the glass-metal interface. The results obtained are summarized in the table, together with results on Nilo K–C.40 glass and Nilo 475–C.12 glass, two commercially-used sealing combinations.

Sealing combination	*No. of seals tested*	*Mean breaking load (lb)*	*Coefficient of variation %*
Zirconium–C.40 glass	8	96·4	6·4
Titanium–C.77 glass	9	109·0	1·3
Titanium–C.78 glass	6	109·6	1·0
Nilo K–C.40 glass	5	85·0	4·8
Nilo 475–C.12 glass	6	69·3	6·5

The zirconium seals usually failed by the glass shearing fairly cleanly away from the metal. With the titanium seals this did not occur; the wire broke near the seal and the glass shattered.

CONCLUSIONS

Satisfactory glass seals can be made to zirconium and titanium.[2] Because of the present high cost of these metals, it is unlikely that they will be extensively used in seal making at the moment, except in applications where special properties are required of the sealing metal, e.g. resistance to chemical attack or lack of ferromagnetism.

ACKNOWLEDGEMENTS

The authors wish to thank Mr. L. J. Davies, Director of Research for permission to publish this paper.

REFERENCES

(1) Redston, G. D., and Stanworth, J. E. *J. Soc. Glass Technol.*, **29**, p. 48 (1945).
(2) British Patent Application No. 33041/53.

Load-deflexion relations of rubber bush mountings

By J. E. Adkins, M.Sc., Ph.D., and A. N. Gent, B.Sc., A.Inst.P., The British Rubber Producers' Research Association, Welwyn Garden City, Herts.

[*Paper received* 25 *January*, 1954]

The relation between applied force and deflexion for bonded cylindrical rubber bush mountings of various dimensions is investigated theoretically and experimentally for four principal modes of deflexion. For symmetrical deformations of the elastic material, a theoretical treatment based on the classical linear theory of elasticity is found to predict adequately the variation of stiffness with bush length. In the case of unsymmetrical deflexions the theoretical treatment is limited to the derivation of approximate values for the limiting stiffnesses for long and short bushes. These theoretical values of limiting stiffness are compared with the values measured experimentally.

1. INTRODUCTION

Rubber bush mountings, consisting of cylindrical rubber tubes bonded on their outer and inner curved surfaces to effectively rigid metal cylinders, are widely used as engineering components. Such applications involve the displacement of the two rigid boundary cylinders relative to each other, often in an unsymmetrical manner. Four principal modes of deflexion may be distinguished, however, which may be produced by fixing the outer cylinder while the inner one undergoes the following displacements:

(i) a rotation about its axis, which we shall refer to as a torsional deflexion,

(ii) a translation in which each point moves parallel to the axis, or axial deflexion,

(iii) a translation in which each point moves through an equal distance perpendicular to a plane containing the axis, or radial deflexion, and

(iv) a rotation of the axis in a radial plane about a point on itself midway between the plane ends of the elastic material, or tilting deflexion.

It is the object of the present work to examine, theoretically and experimentally, the behaviour of cylindrical bushes when deformed as described above.

For practical purposes, the mechanical behaviour of a given bush under each type of deformation may be characterized by its stiffness, which may be defined as the force (or couple) required to produce a unit translational (or angular) displacement. To determine this quantity theoretically, it is necessary to solve the elastic problem in which a cylindrical annulus of finite length is deformed by means of a relative motion of its curved surfaces, while its initially plane ends remain free from applied forces. This presents formidable difficulties, but the two limiting cases where the length of the bush is either very large or very small compared with the difference in radii of the two cylindrical surfaces, are more tractable. In the former, end effects may usually be neglected, while in the latter, the annulus approximates to a disk, and for classically small deformations, the standard methods for the bending of elastic plates may be employed.

From general considerations, we should expect the non-linear stress-strain relationships which apply for large deformations of elastic materials to yield a non-linear variation of displacement with applied force, thus making the measured stiffness dependent upon the magnitude of the deflexion. In practice, however, the rigid boundary surfaces limit the range of deformation obtained, and this, coupled with conditions of symmetry, is usually sufficient to ensure that the classical linear theory of elasticity yields an adequate approximation. The elastic properties of the rubber, considered as an isotropic, incompressible material, may then be completely specified by the modulus of rigidity μ, and Poisson's ratio, which has the value $\frac{1}{2}$.

2. THEORETICAL CONSIDERATIONS

We shall assume the elastic material of the bush to form an annulus of length l, and external and internal radii a and b respectively. We may conveniently specify the configuration of the undeformed body by means of a cylindrical polar coordinate reference frame (r, θ, z) chosen so that the z-axis coincides with the axis of the annulus, and the origin lies midway between its ends. The elastic material is then bounded by the surfaces $z = \pm l/2$, $r = a$, $r = b$. We shall suppose the outer surface to be fixed, and consider the deformations produced by the types of motion of the inner boundary surface described in Section 1.

(i) *Torsional deflexion.* It has been shown by Stevenson[1] and Rivlin[2] that the couple M necessary to rotate the inner cylindrical bounding surface $r = b$ about its axis through an angle ψ is given by

$$M = \frac{4\pi\mu a^2 l}{\chi^2 - 1}\psi \tag{1}$$

where $\chi = a/b$. Rivlin's derivation of equation (1), which is not confined to classically small deformations, assumes a state of plane strain in the elastic material. This can only be maintained in an annulus of finite length by the application of suitable forces to the initially plane ends at $z = \pm l/2$, but since these forces are of the order ψ^2 they can be neglected if ψ is sufficiently small.

(ii) *Axial deflexion.* Rivlin[2] has also considered the case of axial deflexion and has shown that the force L required to displace the inner cylindrical surface $r = b$ through a distance ω_0 (not necessarily classically small) parallel to its axis is given by

$$L = \mu Q_1 l \omega_0 \tag{2}$$

where

$$Q_1 = 2\pi / \log \chi \tag{3}$$

A longitudinally uniform state of simple shear of the elastic material is assumed, a condition which could only be realized in an annulus of finite length by the application of suitable forces to the initially plane ends. Moreover, these forces are now of the order ω_0, so that end effects cannot be made negligible merely by restricting the magnitude of the deformation.

When the length l is small compared with the annular distance $a - b$, the mounting may be regarded as a circular disk and the relations between force and deflexion obtained for small displacements by employing the classical theory for the bending of elastic plates. Assuming that the disk is bent to a state of generalized plane stress,[3] the displacement ω parallel to the z-axis of any point on the middle plane $z = 0$ must satisfy the biharmonic equation

$$\nabla^4\omega \equiv \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)^2 \omega = 0 \tag{4}$$

The resultant force L necessary to produce the deflexion is given by

$$L = -2\pi b[N]_{r=b} \qquad (5)$$

where

$$[N]_{r=b} = -\left[G\frac{\partial}{\partial r}(\nabla^2\omega)\right]_{r=b} \qquad (6)$$

is the stress-resultant acting parallel to the z-axis on the surface $r = b$, and $G = \mu l^3/3$ is the flexural rigidity of the disk. Since the elastic material is bonded to the rigid boundary surfaces we have the conditions

$$\left.\begin{aligned} &\omega = 0, \quad \frac{\partial\omega}{\partial r} = 0 \text{ at } r = a \\ &\omega = \omega_0, \frac{\partial\omega}{\partial r} = 0 \text{ at } r = b \end{aligned}\right\} \qquad (7)$$

A radially symmetrical solution of equation (4), which satisfies conditions (7) is readily obtained, and combination of this with equations (5) and (6) yields

$$L = \mu Q_2 l^3\omega_0/a^2 \qquad (8)$$

where

$$Q_2 = \frac{16\pi\chi^2(\chi^2-1)}{3[(\chi^2-1)^2 - 4\chi^2(\log\chi)^2]} \qquad (9)$$

For bushes of moderate length (i.e. $l \simeq a - b$) neither of the limiting formulae (2) or (8) is adequate, but we may, however, obtain an approximate estimate of the stiffness in axial deflexion by regarding the resultant displacement of the inner cylinder as the sum of separate displacements ω_0', ω_0'' due to shearing and bending respectively of the elastic material. This treatment neglects detailed end effects, and is analogous to that employed by Rivlin and Saunders[4] in deriving an apparent rigidity modulus for cylindrical shear mountings. From formulae (2) and (8) we thus have

$$\omega_0' = L/\mu Q_1 l, \ \omega_0'' = a^2 L/\mu Q_2 l^3 \qquad (10)$$

and since $\omega_0 = \omega_0' + \omega_0''$, we obtain by combining these results

$$L = \mu l\omega_0 \Big/ \left(\frac{1}{Q_1} + \frac{a^2}{l^2 Q_2}\right) \qquad (11)$$

(iii) *Radial deflexion.* The two-dimensional elastic problem in which the inner cylinder undergoes a translation such that each point is displaced through a distance ϵ parallel to a radial line $\theta = 0$, $z = 0$, has been solved, for classically small displacements, by Stevenson.[1] The resultant force which must be applied to the inner cylinder in the direction of motion to produce the displacement ϵ is given by

$$F = \frac{2\pi\mu\kappa(\kappa+1)(\chi^2+1)l\epsilon}{\kappa^2(\chi^2+1)\log\chi - (\chi^2-1)} \qquad (12)$$

where κ is a constant. For an incompressible material in plane strain, we have $\kappa = 1$ and then $F = F_L$ where

$$F_L = \frac{4\pi\mu(\chi^2+1)l\epsilon}{(\chi^2+1)\log\chi - (\chi^2-1)} \qquad (13)$$

To maintain this deformation in an annulus of finite length, normal forces of the order ϵ must be applied to the initially plane ends. If we make the assumption of generalized plane stress, which is appropriate to the limiting case of a circular disk, we have $\kappa = 5/3$ and $F = F_s$, where

$$F_s = \frac{80\pi\mu(\chi^2+1)l\epsilon}{25(\chi^2+1)\log\chi - 9(\chi^2-1)} \qquad (14)$$

We may observe from symmetry considerations that for large deflexions F must be an odd function of ϵ. Terms neglected in employing the linear approximation (12) are thus of the order ϵ^3 and are not likely to become appreciable in practice.

(iv) *Tilting deflexion.* We shall now suppose the inner cylindrical boundary to be displaced so that its axis is rotated by means of a couple of magnitude M_T about its middle point, through a small angle β, in the radial plane $\theta = 0$. The deformation thus produced is neither radially symmetrical, nor longitudinally uniform, and satisfactory analysis becomes inherently more difficult than for the previous cases considered. For this reason, only an approximate treatment of the two limiting cases will be attempted.

In order to obtain an approximation for a sufficiently long annulus, let us consider the deformation of a section of the material bounded initially by planes perpendicular to the z-axis at distances z, $z + dz$ from the origin. If this section is sufficiently far from the origin, the displacement of the inner boundary surface may be regarded approximately as a translation through a distance βz normal to the z-axis, produced by a force dF acting in the direction of motion. The magnitude of this force will depend upon the nature of the deformation in the elastic material, but if we neglect the effect of the free ends and assume a state of plane strain, we may apply equation (13) with l, ϵ replaced by dz, βz respectively. Since the moment of this force about the origin is $dM_T = z\,dF$ we have, neglecting end effects and the effect of forces acting tangential to the inner cylindrical boundary,

$$M_T = \int_{-l/2}^{l/2} \mu\beta R_L z^2 dz = \frac{1}{12}\mu R_L l^3\beta \qquad (15)$$

where

$$R_L = \frac{4\pi(\chi^2+1)}{(\chi^2+1)\log\chi - (\chi^2-1)} \qquad (16)$$

In deriving equation (15), the process of taking moments about the origin has the effect of magnifying errors due to the free ends, and this formula is therefore only likely to be reasonably accurate for bushes in which the length is very great compared with the radii a and b of the curved boundaries.

The other limiting case of a disk bent to a state of generalized plane stress may be considered by a method similar to that employed for axial deflexion. Equation (4) again applies for the displacement ω of any point on the middle plane $z = 0$, but the boundary conditions now become

$$\left.\begin{aligned} &\omega = 0, \ \partial\omega/\partial r = 0 \text{ at } r = a \\ &\omega = \beta b\cos\theta, \ \partial\omega/dr = \beta\cos\theta \text{ at } r = b \end{aligned}\right\} \qquad (17)$$

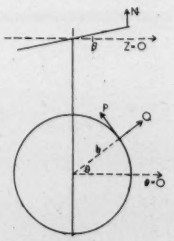


Fig. 1. Tilting deflexion: force and couples on inner boundary

The appropriate solution of equation (4) is therefore

$$\omega = [Ar + Br\log r + (C/r) + Dr^3]\cos\theta \qquad (18)$$

where A, B, C and D are constants which must be chosen so that the conditions (17) are satisfied. In order to obtain the resultant couple M_T, we must consider not only the effect of the stress resultants N given by equation (6), but also that of the flexural and torsional stress-couples P and Q acting upon the surface $r = b$ about lines in the middle plane tangential and normal respectively to that surface (Fig. 1). These are given by[3]

$$\left.\begin{aligned} P &= \tfrac{1}{2}G\Big[-\Big(2\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\Big)\omega \\ &\qquad +\frac{17}{40}l^2\Big(\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\Big)\nabla^2\omega\Big] \\ Q &= \tfrac{1}{2}G\frac{\partial}{\partial r}\Big[\frac{1}{r}\frac{\partial}{\partial\theta}\Big(\omega+\frac{17}{40}l^2\nabla^2\omega\Big)\Big] \end{aligned}\right\} \qquad (19)$$

evaluated at $r = b$. If M_N, M_P, M_Q are the resultant couples due to N, P and Q respectively we readily obtain

$$M_T = M_N + M_P + M_Q \qquad (20)$$

where

$$\left.\begin{aligned} M_N &= -\int_0^{2\pi} Nb^2\cos\theta\, d\theta \\ M_P &= \int_0^{2\pi} Pb\cos\theta\, d\theta \\ M_Q &= \int_0^{2\pi} Qb\sin\theta\, d\theta \end{aligned}\right\} \qquad (21)$$

Combining equation (6) with equations (17) to (21) we have

$$M_T = \frac{4\pi\mu(\chi^2+1)l^3\beta}{3[(\chi^2+1)\log\chi-(\chi^2-1)]} \qquad (22)$$

3. EXPERIMENTAL RESULTS

In order to investigate experimentally the relations derived in Section 2, bush mountings were prepared by a transfer moulding process. The metal parts with the relevant surfaces coated with a suitable bonding cement were placed in the composite mould, followed by injection and subsequent vulcanization of the rubber in the annular space between them. The rubber compound employed contained no additives other than those required for effecting vulcanization. As the thermal expansion coefficient of rubber exceeds that of steel, stresses were induced in the bonded surfaces on cooling from vulcanization temperatures sufficient, in the case of extremely long bushes, to cause separation of the rubber from the inner metal part. These stresses, when insufficiently large to cause bond failure, have been ignored in the following discussion of experimentally observed load-deflexion relations since the deformations associated with them are not considered large enough to affect substantially the elastic behaviour.

To investigate the variation of stiffness with annulus length, eight bushes of different lengths but equal external and internal radii a and b were employed, and measurements made of the load-deflexion relations under torsional, axial, radial and tilting deflexions. Measurements were also made on a single bush having a widely different value for the ratio a/b to verify the predicted dependence of the various stiffnesses on this quantity. Substantially linear, reversible relations were obtained in all cases, confirming the general predictions of Section 1. Values for the slopes of the linear relations obtained, or stiffnesses, are given in Table 1, together with the dimensions of the cylindrical rubber annuli. The maximum linear deflexions imposed were of the order of 0·1 cm during measurements of axial and radial load-deflexion relations, while the maximum angular deflexions imposed in the course of the torsional and tilting measurements were of the order of 0·1 and 0·02 radian respectively.

(i) *Torsional and axial deflexions.* The values given in Table 1 for the rigidity modulus μ of the elastic material of each mounting were obtained by means of equation (1) from the measured values of torsional stiffness, since, in this experiment, errors due to end effects may be made negligibly small by restricting the angle of twist. Reduced values of the axial stiffness per unit length A_1 have been calculated by dividing the measured axial stiffness by the length of the rubber annulus, and by the appropriate value for the rigidity modulus. The values of the reduced axial stiffness per unit length obtained in this way are given in Table 2, together with the corresponding theoretical values A_2 calculated from equation (11). Reasonable agreement is seen to obtain.

(ii) *Radial deflexion.* Reduced values for the radial stiffness per unit length R have been calculated similarly from the measured radial stiffnesses and are given in Table 2, together with the corresponding theoretical values R_L, R_S calculated from equations (13) and (14) respectively. The values of R for bushes 1 to 8 obtained from the experimentally determined radial stiffnesses are plotted against length of annulus in Fig. 2, and it is seen that for small values of l the experimental curve approaches the lower limit R_S for a disk of elastic material. The values of R for the longer

Table 1. *Measured stiffnesses*

Bush no.	Length l (cm)	External radius a (cm)	Internal radius b (cm)	Torsional stiffness (kg.cm/rad)	Axial stiffness (kg/cm)	Radial stiffness (kg/cm)	Tilting stiffness (kg.cm/rad)	Rigidity modulus μ (kg/cm^2)
1	0·635	1·29	0·622	14·75	15·7	51·5	9·30	3·66
2	1·27	1·29	0·622	37·2	45·9	199	40·0	4·62
3	1·905	1·29	0·622	41·5	52·3	258·5	75·5	3·44
4	2·54	1·29	0·622	72·1	98·4	582	224	4·48
5	3·175	1·29	0·622	69·0	89·5	575	345	3·42
6	3·81	1·29	0·622	89·4	124·6	839	760	3·70
7	4·445	1·29	0·622	93·3	123	903	1110	3·31
8	5·08	1·29	0·622	119	154	1119	1880	3·69
9	9·65	3·335	0·715	310	178	620	3960	4·76

Table 2. *Comparison of theoretical and experimental reduced stiffnesses*

Bush no.	A_1 experimental	A_2 calculated from eq. (11)	R experimental	R_S calculated from eq. (14)	R_L calculated from eq. (13)	R_B experimental	T experimental	T_S calculated from eq. (22)	T_L calculated from eq. (15)
1	6·75	6·83	22·1	19·95	117·5	117	9·95	39·27	9·82
2	7·85	8·08	33·9	19·95	117·5	127·5	4·22	39·27	9·82
3	7·97	8·37	39·5	19·95	117·5	123	3·18	39·27	9·82
4	8·65	8·47	51·2	19·95	117·5	107·5	3·05	39·27	9·82
5	8·25	8·52	53·0	19·95	117·5	—	3·15	39·27	9·82
6	8·85	8·55	39·3	19·95	117·5	127·5	3·72	39·27	9·82
7	8·35	8·57	61·5	19·95	117·5	—	3·82	39·27	9·82
8	8·22	8·58	59·8	19·95	117·5	118	3·89	39·27	9·82
9	3·88	4·02	13·5	8·3	20·0	18·5	0·925	6·67	1·667

bushes, however, are seen from Fig. 2 and Table 2 to fall substantially below the theoretical upper limit R_L derived from equation (13). This discrepancy could be ascribed either to the finite lengths of the bushes here employed rendering the assumption of plane strain, from which equation (13) is derived, inadequate, or to appreciable flexibility of the supposedly rigid inner metal part under the deflecting forces applied at its ends.

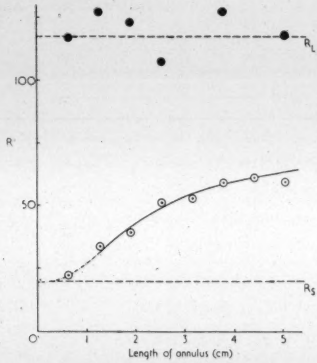


Fig. 2. Effect of length of annulus on the reduced radial stiffness. Values of R and R_B denoted by open and filled circles respectively

Attempts to test the first of these alternative explanations by proceeding to lengths greatly in excess of those already employed would appear to be impracticable since the possibility of appreciable flexure of the inner metal member would thereby be increased. Moreover, the thermal stresses set up on cooling after vulcanization would be correspondingly larger, enhancing the possibility of breakdown of the inner rubber-to-metal bond. The alternative procedure of applying suitable normal forces to the free ends of the elastic material to ensure plane strain was therefore adopted. This was achieved by pressing lubricated metal rings into close contact with the free surfaces to constrain the rubber, sufficient clearance being left between the rings and the inner metal cylinder for small radial deflexions to be imposed. The measured values of reduced radial stiffness, denoted by R_B in Table 2, are seen to be in reasonable agreement in all cases with the theoretical upper limit R_L.

(iii) *Tilting deflexion.* It is seen from equations (15) and (22) that the length l of the rubber annulus enters into the limiting expressions for the tilting stiffness as l^3. The appropriate value of the reduced stiffness per unit length is therefore given by $T = M_T/(\mu\beta l^3)$. We shall denote the theoretical values of T obtained from equations (15) and (22) by T_L and T_S respectively.

The measured values of T are given in Table 2 and, for bushes 1 to 8, are plotted against l in Fig. 3. These experimental values are seen to approach the high theoretical value T_S as l decreases, although the experimental points for the longer bushes lie well below the other limiting value T_L. This discrepancy is not surprising, since in deriving the limit T_L we have considered each section of the bush to be in a state of plane strain, while the experiments on radial deflexion have revealed the inadequacy of this assumption even for the longest of the bushes employed. Moreover, it has already been observed that the method of derivation of equation (15) is likely to accentuate errors arising in this manner. On the other hand, we should expect the limit T_L', obtained by replacing R_L in equation (15) by the lower limit R_S obtained from equation (14) on the assumption of generalized plane stress, to lie below the experimental values of T, and it may be seen from Fig. 3 that this conclusion is borne out in practice.

In view of the inadequacy of the plane strain assumption for bushes with free ends, measurements were made of the

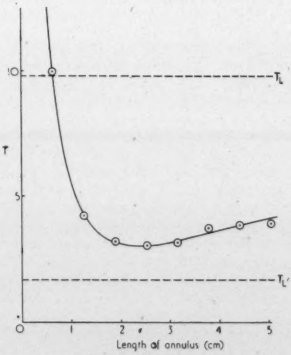


Fig. 3. Effect of length of annulus on the reduced tilting stiffness

tilting stiffness of some of the longer bushes while lubricated metal rings were pressed into close contact with the free rubber surfaces. This procedure is clearly less satisfactory than in the case of radial deflexion, since the effect of small frictional forces at the plane ends will be greater. In practice, although the experimental scatter was considerable, the measured values of T were found to lie much closer to, but somewhat below the theoretical limit T_L.

4. CONCLUSIONS

It has been confirmed experimentally that a general indication of the load-deformation behaviour of cylindrical rubber bush mountings may be obtained from an approximate theoretical treatment based on the classical linear theory of elasticity. Experiments on torsional deflexion may be employed to obtain a value for the modulus of rigidity for the elastic material, and when this value is introduced into the formula for the axial stiffness, satisfactory agreement is obtained between the theoretical and experimental results over the entire range of lengths investigated. For the less symmetrical radial and tilting deflexions, attention has been confined to the derivation of the limiting values of the stiffness for very long and very short bushes. In both cases, the experimental values approach the theoretical limit for short bushes, but in neither case is the limit for long bushes attained within the range of lengths employed. This suggests that the assumption of plane strain on which the theoretical formulae are based is only likely to be admissible for bushes much longer than those commonly encountered.

The present work has been concerned mainly with the variation of stiffness with length, and eight bushes of equal radial dimensions have been employed. From the theoretical expressions, and also from physical considerations, it is evident that the forces required to produce a given deflexion will vary greatly with the ratio χ of the external and internal radii of the cylindrical annulus. Measurements on a ninth bush of widely differing radial dimensions suggest, however, that the general conclusions reached from the previous experiments will be generally applicable.

ACKNOWLEDGEMENTS

One of the authors (J. E. A.) wishes to thank the Director of the Davy Faraday Laboratory of the Royal Institution for the provision of facilities during the progress of this work, which forms part of the programme of research undertaken by the Board of the British Rubber Producers' Research Association.

REFERENCES

(1) STEVENSON, A. C. *Phil. Mag.*, **34**, p. 766 (1943).
(2) RIVLIN, R. S. *Phil. Trans. A.*, **242**, p. 173 (1949).
(3) LOVE, A. E. H. *The Mathematical Theory of Elasticity*, 2nd Ed., p. 452 (Cambridge: University Press, 1906).
(4) RIVLIN, R. S., and SAUNDERS, D. W. *Trans Inst. Rubber Industr.*, **24**, p. 296 (1949).

The experimental determination of the speed of a vacuum pump and of components of a vacuum system

By C. W. OATLEY, M.A., M.Sc., Engineering Laboratory, University of Cambridge

[*Paper first received* 30 *March, and in final form* 29 *April*, 1954]

A description is given of a method of measuring the speed of a vacuum pump which appears to have many advantages over existing methods. It can be used at pressures down to at least 10^{-4} mm and is independent of the absolute accuracy of the calibration of the gauge used to measure these pressures.

By an extension of the method it becomes possible to measure the resistance to the flow of rarefied gas of any tube or other channel of which a small-scale model can be made. Results are given for right-angle bends, T-pieces and other components which are likely to be used in the construction of a complete vacuum system.

1. INTRODUCTION

Several methods have been devised for measuring the speed of a vacuum pump,[1] but in most of them a knowledge of the absolute calibration of a low-pressure gauge is directly involved. Although, in principle, the calibration of such a gauge for a particular gas presents no difficulty, appreciable errors may arise in practice.

In the method to be described, the determination of the speed of a pump involves only the ratios of pressures which are of the same order of magnitude and these ratios can conveniently be measured with an ionization gauge which, for this purpose, need not be calibrated so long as it is working within its linear range. The gauge must indeed be calibrated to find the pressure at which the speed has been measured, but, since speed generally varies slowly with pressure, this calibration need not be very accurate.

It is assumed throughout that the pump is working at a sufficiently low pressure for collisions between molecules to be neglected. At such pressures it is convenient to define the conductance S_1 of a component of a vacuum system (e.g. a connecting tube or a baffle chamber) as the nett volume of gas, measured at the pressure of entry, which would flow into the component in unit time if one end were connected to a gas reservoir at constant pressure p_1 while the other was maintained at zero pressure. S_1 depends on the molecular weight and the temperature of the gas but, so long as inter-molecular collisions are negligible, it is independent of p_1.

Calculations of values of S_1 for an aperture in a thin plane plate and for a cylindrical tube, the length of which is very large compared with its radius, were first given by Knudsen[2] and these problems have also been discussed by Smoluchowski.[3] The general case of a cylindrical tube of arbitrary length and radius has been considered by Clausing,[4] but the more complicated shapes assumed by components of a practical vacuum system are hardly amenable to calculation and very few experimental measurements have been published for them. It is shown below that the apparatus, developed originally for measuring the speed of a pump, can also be used to determine the speed of a component of any

shape for which a small-scale model can be constructed, and results are given for a number of such components.

2. THE MEASUREMENT OF PUMP SPEED

The general arrangement of apparatus is shown in Fig. 1(*a*). The pump to be tested is connected to the brass tube *A* which is sufficiently wide and short to have negligible effect on the speed of pumping. Tube *A* is connected to a glass vessel *B* through a circular aperture *C* which is one of six, of different sizes, cut in a circular brass plate *D*. This plate is free to rotate about a spindle *E* but a spring washer keeps it pressed against the base plate *F*. The two surfaces which are in contact are ground flat and coated with a film of vacuum

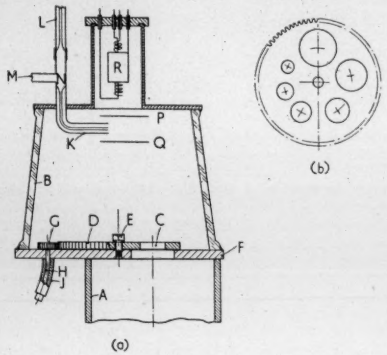


Fig. 1. (*a*) General arrangement of apparatus for measuring the speed of a vacuum pump. (*b*) Arrangement of holes *C* in plate *D*

grease to form a reasonably good seal. Teeth are cut round the periphery of plate *D* so that it can be turned by a pinion wheel *G* which is actuated from outside the vacuum by means of the bent shaft *H*, surrounded by a short length of pressure tubing *J* which is closed at the bottom end. Thus any of the apertures in plate *D* can be brought into position between the pump and vessel *B*. The arrangement of holes in plate *D* is shown in Fig. 1(*b*). Two interchangeable plates are provided and the hole diameters for each are given in Table 1. The plates are about 12 cm in diameter and 0·5 cm thick.

Air leaks into the vessel through the glass capillary tube *K* and it is essential that the rate of leak should remain constant. It might well be possible to obtain sufficient constancy with a needle-valve or similar device connected directly to the atmosphere, but the arrangement used in the present experiments was thought likely to be more reliable. A second capillary tube *L*, open to the atmosphere at one end, is connected as shown and tube *M* is joined to a rotary oil-pump so that the pressure at the point *N* is of the order of 1 mm of mercury. The magnitude of this pressure and hence the rate of leak of air into the main vessel can be controlled by choice of the length and internal diameter of the tube *L*. Suitable rates can be obtained with tubes about 10 cm long and a few tenths of a millimetre in diameter.

Inside the main vessel the end of the tube *K* passes into a bundle of glass wool which is held in place by two nickel disks *P* and *Q*, the whole being attached to *K* by wire. The incoming gas is thus caused to diffuse laterally towards the side of the vessel *B* and there is no tendency for it to form a molecular beam either towards the hole *C* or into the ionization gauge *R* which is located at the top of the vessel.

Let S_1 be the speed of the pump at a particular pressure and S_2 the conductance of the hole *C* which joins the main vessel to the pump. S_2 can be accurately calculated from data given by Clausing.[4] Then the total speed S of the pump in series with the hole is given by

$$\frac{1}{S} = \frac{1}{S_1} + \frac{1}{S_2} \tag{1}$$

The mass of gas withdrawn per second from the main vessel is equal to kpS where k is a constant and p is the pressure in the main vessel. Under equilibrium conditions this must be equal to the constant mass m of gas entering the vessel per second through the leak.
Thus

$$kpS = m$$

and, from equation (1),

$$\frac{k}{m}p = \frac{1}{S_1} + \frac{1}{S_2} \tag{2}$$

If, therefore, for different sizes of the hole *C* but with all other conditions kept constant, p is plotted against $1/S_2$, the resulting curve will be a straight line with intercept equal to $1/S_1$.

Since the determination of S_1 is independent of the slope of the curve, any quantity which is proportional to p may be plotted instead of p itself. Hence the readings of an uncalibrated ionization gauge may be used so long as the pressure is sufficiently low for the gauge to be working in its linear range. Again, if the gas emerging from the glass wool between the plates *P* and *Q* has a tendency to move in a particular direction, so that the pressure in the ionization gauge is not quite the same as that immediately above the aperture *C*, this will not matter provided the ratio of the two pressures is constant. Similarly, the measurement of the speed of the pump will be unaffected by small unintended leaks if these cause air to enter the apparatus above the aperture *C*.

The hot filament of the ionization gauge will remove a certain amount of gas by chemical combination and this may affect the results in two ways. In the first place the total mass of gas flowing per second into the pump will be reduced. It is easy to show, however, that under the normal conditions of operation, the rate of removal of gas by the gauge is negligible compared with the rate of removal by the pump, so this effect may be disregarded. A more serious possibility is that the action of the filament may cause the pressure in the gauge to be less than that in the main vessel. As has already been shown, this is immaterial so long as the ratio of the two pressures is constant and this will be the case if the rate of removal of gas by the filament is proportional to the pressure in the gauge. The rate of removal will presumably depend directly on the rate at which molecules strike the filament and this is proportional to the pressure in the gauge. Hence the pumping action of the gauge should not affect the final results.

At first sight it might appear that varying the size of the aperture *C* during a run would necessarily change the pressure immediately below this aperture, and hence that the method would only be valid for a pump, the speed of which was independent of pressure. However, this is not the case; gas enters the apparatus at a constant rate and, for equilibrium, it must leave at the same rate. Thus the pressure in the

throat of the pump remains constant and the pressure difference between the two sides of C adjusts itself so that gas flows through C at the required rate. With the former notation, the pressure p_1 in the throat of the pump would be equal to p if S_2 were infinite. Hence p_1 can be determined from the plot of p against $1/S_2$, since it is that value of p for which $1/S_2$ is zero.

3. TYPICAL RESULTS FOR THE MEASUREMENT OF PUMP SPEED

To obtain accurate results, the sizes of the holes in plate D must be chosen to suit the speed of the pump under test. The sizes given for plate 1 in Table 1 cover the range from about 10 to 80 l./s satisfactorily and can be used with diminishing accuracy outside this range. The diameters of the individual holes have been chosen to give uniformly spaced points on the plot of p against $1/S_2$.

Table 1 also gives the effective area of each hole, that is, the area of a hole in an infinitely thin plate which would have the same conductance as the actual hole. These effective areas are calculated from Clausing's results and, as is shown in Section 4, the conductance of a hole in litres per second for air at room temperature is equal to 11·7 times its effective area.

Table 1. *Hole sizes for plate D (Fig. 1)*

	Hole no.	1	2	3	4	5	6
Plate 1	*Hole radius* (cm)	1·770	1·240	1·015	0·875	0·778	0·716
	Effective area (cm²)	8·47	3·93	2·53	1·81	1·39	1·15
Plate 2	*Hole radius* (cm)	1·041	0·793	0·653	0·564	0·507	0·468
	Effective area (cm²)	2·676	1·456	0·937	0·667	0·521	0·431

In Fig. 2 are shown curves for two sets of results on the same pump, working with different input pressures. These curves are quite typical of many that have been obtained

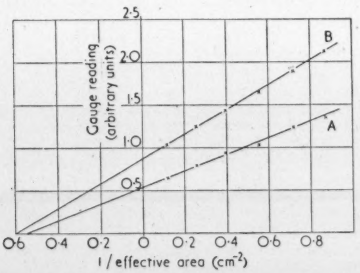


Fig. 2. Curves for the measurement of pump speed

with a variety of pumps. For curve A the input pressure was $7{\cdot}4 \times 10^{-4}$ and for curve B $1{\cdot}2 \times 10^{-4}$ mm of mercury. The corresponding speeds are 21·3 and 19·8 l./s respectively.

When measurements at relatively low pressures are being made, gas given off from the walls of the apparatus may contribute significantly to the total pressure registered by the ionization gauge. If this gas were emitted at a constant rate, it would not affect the final result, but should it contain a component such as water vapour, which is readily absorbed, this may not be the case. The procedure adopted to deal with this situation has been as follows. Immediately after the normal set of readings has been taken, the air leak is cut off and, when equilibrium has been established, a new set of pressure readings is taken. These readings are then subtracted from the first set and the differences taken to be independent of gas given off from the walls. This assumption is probably not far from the truth and the procedure appears to be satisfactory for pressures down to at least 10^{-4} mm.

The consistency of the readings in a given set suggests that the value obtained for the speed of the pump should be accurate to within about five per cent.

4. THE FLOW OF RAREFIED GAS THROUGH CYLINDRICAL TUBES OF FINITE LENGTH

The apparatus described above can readily be adapted to the measurement of the rate of flow of rarefied gas through a tube of any shape or size. For such a tube XY let the end X be connected to a reservoir of gas at constant low pressure, while the end Y is joined to a pump of infinite speed, so that any gas molecule which reaches Y is removed by the pump. Of the molecules which enter X some pass through the tube to Y and the pump while others, after collision with the walls, return through X to the reservoir.

Let W be the average probability that a molecule, which passes from the reservoir through the entrance aperture at X, will succeed in reaching Y before being returned to the reservoir. The pressure throughout the tube is assumed to be so low that collisions between molecules are of negligible importance; thus the chance of any one molecule getting through the tube is unaffected by the presence of the others. Then if n is the number of molecules per unit volume of the reservoir, c their mean velocity and a the area of the entrance aperture of the tube, the number of molecules falling per second on the entrance aperture is, from kinetic theory, $nca/4$ and the number which leave the reservoir permanently per second is $ncaW/4$. Hence the volume of gas withdrawn per second from the reservoir, measured at the pressure existing in the reservoir, is $caW/4$ and this is the quantity generally known as the conductance of the tube. When the molecular weight and temperature of the gas are known, $c/4$ can be calculated; for air at room temperature it has a value which may be written in the form 11·7 l./s.cm².

For a molecule entering the aperture at X at a particular point and travelling in a particular direction with respect to the tube, the chance of reaching Y would be unchanged if all the dimensions of the tube were scaled in the same ratio. Hence W is the same for all tubes which are geometrically similar and, in consequence, is a convenient parameter to tabulate. Furthermore, for the experimental determination of W, a model of any convenient scale may be used.

Suppose such a model to be constructed of a size to fit closely into No. 1 hole of the plate D in Fig. 1, the bottom end of the tube falling flush with the underside of the plate. With a constant rate of leak of air into the top part of the apparatus, a plot of ion-gauge readings against the reciprocal of effective hole area will yield a straight line graph for holes Nos. 2 to 6 inclusive (Fig. 3). Using this graph and the ion-gauge reading for hole No. 1 (with the model in place), the effective area for the model can be determined. Then, knowing the actual area of the entrance aperture of the model, the value of W for the model can be calculated. For measurements of this kind the hole sizes given for plate 2 in Table 1 have been found convenient.

To check the accuracy of this method, measurements were made on a series of cylindrical tubes of the same radius r but of different lengths l. Theoretical calculations of W for such tubes have been made by Clausing[4] and his values,

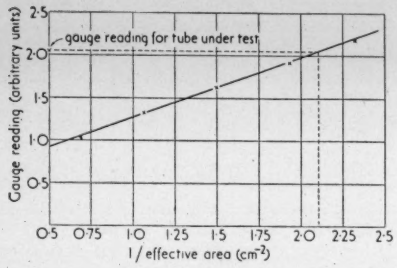


Fig. 3. Curve for the measurement of tube speed

together with those obtained experimentally, are given in Table 2. The results given in column (*a*) were obtained using pressures about ten times as great as were used for column (*b*).

Table 2. *Theoretical and experimental values of W*

l/r	W (Clausing)	Experimental values of W (*a*)	(*b*)
2	0·518	0·487	0·516
4	0·358	0·359	0·341
6	0·280	0·251	0·251
8	0·231	0·228	0·214
10	0·196	0·178	0·173

The two sets of experimental values agree with each other rather better than with Clausing's values and they are consistently lower than the latter. The error resulting from slight drift of the experimental conditions was about five per cent and it therefore seems probable that a systematic error of about the same size was also present. This could have been caused by a small variation of pressure across the brass tube A (Fig. 1). When the upper end of A is terminated by a relatively thin plate with a hole in it, molecules from A travelling upwards and falling on the aperture from all directions have a high probability of getting through into the vessel B. If, however, a vertical cylindrical tube is inserted in the hole, molecules from A falling on the aperture in directions nearly parallel to the axis of the tube have a much higher probability of getting through than those which fall at oblique angles. Thus if the pressure in the centre of A were slightly higher than that round the periphery, this would have a greater effect on the gas getting back into B through a tube than through a hole in a thin plate. Since the values of W for tubes were determined by comparison with known values for holes in a plate, the measured values for tubes would be in error by amounts which would increase with the ratio of length to radius.

If this is the correct explanation, the difficulty could be overcome by using a much larger tube A and by inserting baffles between the bottom end of this tube and the pump. However, it was not considered worth while to make this modification, since the agreement between measured and calculated values of W was adequate for the purpose in hand.

5. DETERMINATION OF W FOR CHANNELS OF OTHER SHAPES

The method described above has been used to measure W for a number of channels which are likely to occur as components of a vacuum system. The shapes of these channels are indicated in Fig. 4 and the results obtained are as follows.

(*a*) *Tubes of rectangular cross-section.* All three of the tubes measured had a length of 1·75 cm and a cross-sectional area of 0·6 cm^2. For different ratios b/a [Fig. 4(*a*)], the values of W were

b/a	1	2	4
W	0·39	0·37	0·36

It is interesting to note that, for a circular tube of the same length and area of cross-section, Clausing's value of W is 0·359.

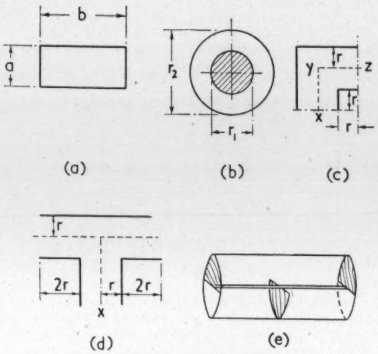


Fig. 4. Details of components

(*b*) *The annular channel between two concentric cylinders.* The outer tube in each case had a diameter of 1·90 cm and was 3·83 cm long. Different values of r_2/r_1 [Fig. 4(*b*)] were obtained by using different inner rods, which were also 3·83 cm long. Each rod in turn was supported from the outer cylinder by a thin pin which could not have had much effect on the flow of gas. Table 3 gives, for each value of r_2/r_1, the area of the annular space, the measured value of W, and Clausing's value of W for a cylindrical tube of the same length and area of cross-section.

Table 3. *Values of W for concentric cylinders*

r_2/r_1	Area of annular space	W (measured)	W (Clausing) for equivalent cylindrical tube
3·02	2·54	0·263	0·347
2·40	2·36	0·244	0·340
2·00	2·14	0·227	0·330
1·71	1·88	0·204	0·317
1·50	1·58	0·184	0·301

(*c*) *Angles and T-pieces.* For a right-angle bend of circular cross-section, of the form shown in Fig. 4(*c*) the measured value of W was 0·30. For such a bend, the mean length of path $x\ y\ z$ along the axes of the two tubes is four times the

radius, and Clausing's value of W, for a straight cylindrical tube with $l/r = 4$, is 0·359. The effect of bending the tube is therefore not very great.

With the T-piece of Fig. 4(*d*), also of circular cross-section, for gas entering through x and leaving by either of the other apertures, the measured value of W is 0·30.

(*d*) *Traps for oil vapour.* A commonly-used form of vapour trap is shown in Fig. 4(*e*). Two models of this type, each containing semi-circular baffles at the two ends and one in the centre, were measured. For a trap length equal to the diameter of the outer tube, the value of W was 0·065. With a trap length twice the tube diameter, W was 0·078. In both cases these values are calculated for an entrance aperture equal to half of the full cross-sectional area of the tube.

ACKNOWLEDGEMENT

The author is indebted to Mr. L. R. Peters who constructed most of the apparatus and assisted him with the experimental work.

REFERENCES

(1) See, for example, DUSHMAN, S. *Scientific Foundations of Vacuum Technique* (New York: John Wiley and Sons, 1949).
(2) KNUDSEN, M. *Ann. Phys.* [*Leipzig*], **28**, pp. 75 and 999 (1909).
(3) SMOLUCHOWSKI, M. v. *Ann. Phys.* [*Leipzig*], **33**, p. 1559 (1910).
(4) CLAUSING, P. *Ann. Phys.* [*Leipzig*], **12**, p. 961 (1932).

Resistance-network analogues with unequal meshes or subdivided meshes

By G. LIEBMANN, D.Phil., F.Inst.P., Research Laboratory, Associated Electrical Industries Limited, Aldermaston, Berks.

[*Paper first received* 1 *February, and in final form* 24 *March*, 1954]

A general method for deriving the finite difference approximations to the partial differential equation div K grad $U = g$, which is solved by resistance-network analogues, is applied to find the resistance values and the currents which have to be fed in for (x, y)- and (r, z)-networks with unequal mesh sizes. Two ways of subdividing a network into finer meshes in the ratio 1 : 2 are then given, and a network subdivision in the ratio 2 : 5 is described. Experimental tests have shown that these subdivisions introduce negligible errors into the measured field distributions.

The resistance-network analogues[1–3] for solving, by electrical measurements, field equations of the type

$$\text{div } K \text{ grad } U = g \qquad (1)$$

where K and g are known functions of the co-ordinates, and U is the unknown function, with prescribed boundary values or boundary gradients, usually employ uniform mesh size. As in the numerical methods for solving equations of the type of equation (1), so in the network analogue technique it is often desirable to work with unequal meshes or locally subdivided meshes, to obtain higher accuracy where the geometry of the problem is complicated, or where the field distribution changes rapidly. This question has already been discussed briefly in the earlier literature,[2–5] but these earlier treatments were incidental and incomplete, and led in some instances to insufficiently accurate network design principles. In this paper, a more complete and more precise treatment of the question will be given.

UNEQUAL MESHES

The resistance-network analogue does not solve the partial differential equation (1), but its equivalent finite difference approximation. In this finite difference method, common to well-known numerical methods and the network analogue method, the area, or volume, within the boundary is divided into a number of meshes, or cells, and the discrete values of U and g at the centres of each mesh or cell are taken as representing all values of U or g within the mesh or cell. The required finite difference approximation to equation (1) at the centre point P_0 of a cell can be derived in its most general form by integrating both sides of equation (1) over the volume Δv_0 of the cell surrounding the point P_0 (see Fig. 1), and then replacing, by Gauss's theorem, the volume integral on the left-hand side by a surface integral. Substituting for the surface integral a sum over the surface elements δA_ν of the cell gives

$$\sum_\nu K_\nu (\partial U/\partial n)_\nu \delta A_\nu = g_0 \Delta v_0 \qquad (2)$$

where K_ν is the value of K at the surface element δA_ν, $(\partial U/\partial n)_\nu$ the normal derivative of U at the surface element δA_ν, and g_0 the value of the function g at the point P_0. Finally, approxi-

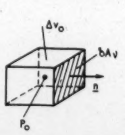


Fig. 1. Volume element Δv_0 surrounding point P_0.

mating the normal derivatives $(\partial U/\partial n)_\nu$ by their difference quotients $(\delta U/\delta n)_\nu$ yields the required finite difference equation, which is then solved by the analogue. This way of deriving the difference equation holds, whatever system of co-ordinates is represented by the analogue apparatus, and whatever the size of Δv or δA_ν, i.e. whatever the mesh size employed. The two approximations involved in this procedure are the assumption that $g = g_0$ over the whole volume Δv_0, which may only be correct if Δv_0 is very small, and that $(\partial U/\partial n)_\nu$ is constant over the surface element δA_ν, and is varying at such a small rate with n that the replacement of $(\partial U/\partial n)_\nu$ by the difference of the values of U at adjacent mesh points divided by the distances of the points is permissible. The general procedure can be readily adapted to any special case by inserting into equation (2) the values of K_ν, $(\partial U/\partial n)_\nu$, δA_ν and Δv_0 for the used system of co-ordinates, and their subdivision into finite steps (meshes) which has been prescribed.

This will be illustrated by deriving the resistance-network

relations for an (x, y)-network representing unequal meshes as shown in Fig. 2. Here $\delta z = \text{const} = 1$, say; let the distances of the point P_0 from its neighbours $P_1 \ldots P_4$ be δx_1,

Fig. 2. Neighbouring points in (x, y)-network with unequal meshes

δy_2, δx_3 and δy_4. Then, for example, for $\nu = 1$, with $K(P_0)$ representing the value of K at P_0, etc.

$$K_1 \simeq [K(P_0) + K(P_1)]/2$$

$$\delta A_1 = [(\delta y_2 + \delta y_4)/2]\delta z = (\delta y_2 + \delta y_4)/2$$

$$(\partial U/\partial n)_1 \simeq (U_0 - U_1)/(-\delta x_1)$$

with corresponding values for $\nu = 2, 3, 4$, and

$$\Delta v_0 = \frac{(\delta x_1 + \delta x_3)}{2} \times \frac{(\delta y_2 + \delta y_4)}{2}\delta z$$

$$= \frac{(\delta x_1 + \delta x_3)(\delta y_2 + \delta y_4)}{4}$$

Inserting this into equation (2), one obtains the required difference expression

$$K_1\frac{(\delta y_2 + \delta y_4)}{2\delta x_1}(U_1 - U_0) + K_2\frac{(\delta x_1 + \delta x_3)}{2\delta y_2}(U_2 - U_0)$$

$$+ K_3\frac{(\delta y_2 + \delta y_4)}{2\delta x_3}(U_3 - U_0) + K_4\frac{(\delta x_1 + \delta x_3)}{2\delta y_4}(U_4 - U_0)$$

$$= \frac{(\delta x_1 + \delta x_3)(\delta y_2 + \delta y_4)}{4}g_0 \qquad (3)$$

This equation, with U represented by the network voltages, is satisfied by the four currents $i_1 \ldots i_4$ flowing in the "net-

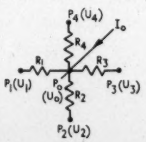


Fig. 3. Resistance-network "star" corresponding to mesh point arrangement of Fig. 2

work star," Fig. 3, with the current I_0 fed in from an external source, if

$$\left.\begin{aligned} R_1 &= [2\delta x_1/(\delta y_2 + \delta y_4)K_1]R_N \\ R_2 &= [2\delta y_2/(\delta x_1 + \delta x_3)K_2]R_N \\ R_3 &= [2\delta x_3/(\delta y_2 + \delta y_4)K_3]R_N \\ R_4 &= [2\delta y_4/(\delta x_1 + \delta x_3)K_4]R_N \end{aligned}\right\} \qquad (4)$$

R_N being a suitably chosen value of resistance (an apparatus constant), and

$$I_0 = -[(\delta x_1 + \delta x_3)(\delta y_2 + \delta y_4)/4R_N]g_0 \qquad (5)$$

As a special case of this consider rectangular meshes as shown in Fig. 4. Then $\delta x_1 = \delta x_3 = h'$, $\delta y_2 = \delta y_4 = h$. Assume, for simplicity, $K_1 = \ldots K_4 = 1$. Then

$$R_1 = R_3 = (h'/h)R_N, \; R_2 = R_4 = (h/h')R_N$$

and

$$I_0 = -(hh'/R_N)g_0.$$

This type of mesh is very useful if one wishes to extend the effective dimensions of a network in one direction only; it is being used by the author, for example, to remove the boundaries of the network to a long distance away from the "working part" of the resistance-network analogue. Of

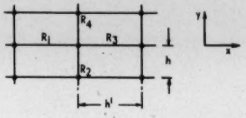



Fig. 4. Mesh points in net with unequal meshes

course, in the analogue apparatus, a "long" mesh may be of the same physical size as a "square" one, as only the *value* of the resistances is changed, and not their size. At the place where the transition from the square to the rectangular mesh occurs in Fig. 5, equations (4) and (5) give these values: $R_1 = R_N$ (the standard value for all square meshes to the left of the join), $R_2 = R_4 = 2R_N/(1 + h'/h)$, $R_3 = (h'/h)R_N$, and $I_0 = (h^2/2R_N)(1 + h'/h)g_0$. The data

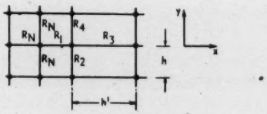


Fig. 5. Mesh junction between square meshes and rectangular meshes

for the network *termination* at an open boundary passing through the point P_0 are obtained at once as a special case by putting $h' = 0$: $R_2 = R_4 = 2R_N$ and $I_0 = h^2g_0/2R_N$.

Similarly, one can derive from equation (2) the design of a network with unequal meshes for solving *axially symmetrical* problems. The "unit cell" is then a cylindrical shell (see Fig. 6) of radii $r_1' = r_0 - (\delta r_1/2)$ and $r_2' = r_0 + (\delta r_2/2)$, and of length $\frac{1}{2}(\delta z_1 + \delta z_2)$ if δz_1 and δz_2 are the distances of the

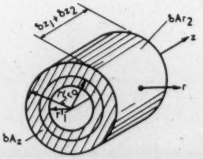


Fig. 6. Volume element in axially symmetrical problems

neighbouring points to P_0 in the axial direction, and $r_1 = r_0 - \delta r_1$ and $r_2 = r_0 + \delta r_2$, the off-axis distances of the neighbours of P_0 in the radial direction, "mesh points" being the intersections of concentric circles with a meridional plane through the cylinder.

Then $\delta A_1 = \delta A_{r_1} = \pi(\delta z_1 + \delta z_2)(r_0 - \delta r_1/2)$,

$$\delta A_2 = \delta A_4 = \delta A_z = \pi[(r_0 + \delta r_2/2)^2 - (r_0 - \delta r_2/2)^2]$$

$$= \pi(r_2 - r_1)(r_1 + 2r_0 + r_2)/4$$

$$\delta A_3 = \delta A_{r_2} = \pi(\delta z_1 + \delta z_2)(r_0 + \delta r_2/2),$$

and

$$\Delta v_0 = \tfrac{1}{2}(\delta z_1 + \delta z_2)\delta A_2.$$

Inserting these values into equation (2), and using for $(\partial U/\partial n) = (\partial U/\partial r)$ the approximations $(U_1 - U_0)/\delta r_1$ and $(U_3 - U_0)/\delta r_2$, leads directly to a generalization, for unequal mesh sizes and different values of K_ν, of the standard difference equations of the usual numerical methods. Within a few mesh lengths from the axis, however, the change of $\partial U/\partial r$ with r is no longer negligible even in the first approximation, and a better first approximation is obtained by observing that in the basic equation (2) the terms $r(\partial U/\partial r)$ play the same part as the $(\partial U/\partial z)$ terms. In this better first approximation, $(r_0 - \delta r_1/2)(U_1 - U_0)/\delta r_1$ becomes replaced by $(U_1 - U_0)/\log_e(r_0/r_1)$, etc. The generalized design formulae for an (r, z)-resistance-network analogue with unequal meshes become therefore:

$$\left.\begin{aligned} R_1 &= \frac{2R_N \log_e (r_0/r_1)}{(\delta z_1 + \delta z_2)K_1} \\ R_2 &= \frac{8\delta z_1 R_N}{(\delta r_1 + \delta r_2)(4r_0 - \delta r_1 + \delta r_2)K_2} \\ R_3 &= \frac{2R_N \log_e (r_2/r_0)}{(\delta z_1 + \delta z_2)K_3} \\ R_4 &= \frac{8\delta z_2 R_N}{(\delta r_1 + \delta r_2)(4r_0 - \delta r_1 + \delta r_2)K_4} \end{aligned}\right\} \quad (6)$$

and $$I_0 = -\frac{(\delta z_1 + \delta z_2)(\delta r_1 + \delta r_2)(4r_0 - \delta r_1 + \delta r_2)}{16R_N} g_0 \quad (7)$$

For a mesh point P_0 on the axis, equations (6) and (7) do not hold but the required network data can be derived as before from equation (2), taking as "unit cell" the cylindrical element of length $(\delta z_1 + \delta z_2)/2$ and radius $\delta r_2/2$. The resistance R_1 does then no longer appear in the axial "resistance star," and the values of the other three resistances become:

$$\left.\begin{aligned} R_2 &= \frac{8\delta z_1 R_N}{(\delta r_2)^2 K_2} \\ R_3 &= \frac{4R_N}{(\delta z_1 + \delta z_2)K_3} \\ R_4 &= \frac{8z_2 R_N}{(\delta r_2)^2 K_4} \end{aligned}\right\} \quad (6a)$$

with $$I_0 = -\frac{(\delta r_2)^2(\delta z_1 + \delta z_2)}{16R_N} g_0 \quad (7a)$$

It is seen that the resistances along the axis remain finite. The design formulae (6), (7) and (6*a*), (7*a*) can, of course, be used to derive the data for rectangular meshes and for the network termination at open boundaries. For $\delta z_1 = \delta z_2 = \delta z$, and $\delta r_1 = \delta r_2 = \delta r$, with $K_1 = \ldots K_4 = 1$, they reduce to the formulae given by the author in an earlier paper.[3]

GEOMETRICAL INTERPRETATION OF UNEQUAL MESHES

A useful geometrical interpretation of the derived design formulae follows immediately from equation (2). It will be stated with reference to the formulae for the (x, y)-network, equations (4) and (5), but applies in suitably modified form to networks based on other co-ordinate systems. Referring to Figs. 7(*a–c*), the *resistance* R_1 specified by equation (4) is equal to the resistance of a strip of material of unit thickness $(\delta z = 1)$, and of specific resistance R_N/K_1, of width $(\delta y_2 + \delta y_4)/2$ and of length δx_1; similarly for resistances R_2, R_3 and R_4. The *current* I_0 fed in from an external source at the network point P_0 is seen to be equal, apart from the multiplying factor g_0/R_N, to the *volume* $(\delta x_1 + \delta x_3)(\delta y_2 + \delta y_4)/4$ of the rectangular prism with sides $(\delta x_1 + \delta x_3)/2$, $(\delta y_2 + \delta y_4)/2$ and $\delta z = 1$ enclosing the network point P_0. This geometrical interpretation of the design formulae is the basis of the rules

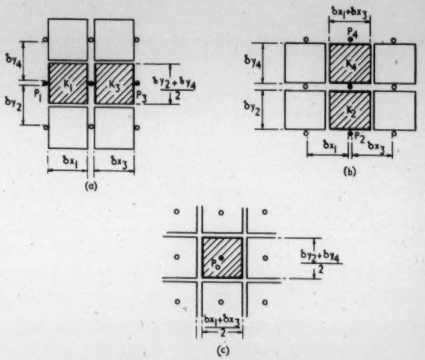


Fig. 7. Illustration of the geometrical interpretation of equations (4) and (5)

which were discussed in some detail in earlier papers,[3,6] for locally modifying a network with square or rectangular meshes to represent curved boundaries; on the basis of the geometrical interpretation one can mostly write down at once the network modifications required in such cases.

MESH SUBDIVISION

The local subdivision of meshes in a resistance-network analogue will be discussed with reference to a square arrangement of meshes in an (x, y)-network, and a value of $K_1 = \ldots K_4 = 1$, to simplify the presentation; the extension to conditions where unequal meshes or different values of K are used, is obtained by combining the features of subdivided nets with those of unequal meshes.

The use of subdivided nets was introduced into the relaxation method of solving numerically the problems described by equation (1) by Allen, Southwell and Vaisey[7] in 1945. In their method, the "coarse" net of square meshes of mesh length h is subdivided partly into square meshes of length $h/2$, this "fine" net being joined to the coarse one by a series of diagonal meshes of mesh length $h/\sqrt{2}$. Part of the fine mesh may then be subdivided further in the same manner. A particular feature of this method of mesh subdivision, as pointed out by the authors, is the necessary assumption that certain mesh points in the joining strip of diagonal meshes are influenced by the values of U at the coarse mesh points, but do not themselves influence the values in the coarse mesh, i.e there is a "one-way traffic" of residuals at certain of the intermediate mesh points. This freedom of the computor to ignore at will the influence on the field distribution of the field values at certain mesh points is not shared by the resistance-network analogue, where the interconnexion of the resistances leads necessarily to an automatic mutual influence of the field values at neighbouring points; indeed, it is this very feature which leads to one of the most

attractive properties of resistance networks, the statistical cancellation of the effect of resistance tolerance errors.[3] The difficulty mentioned can be overcome, when applying Allen, Southwell and Vaisey's method of subdivision to resistance networks, by adding further diagonal resistances to the joining strip of diagonal meshes, but the intermediate strip which "matches" the coarse network into the fine network then becomes complicated, and detailed analysis has shown that it is not better in accuracy than the two simple methods of resistance-network subdivision which will now be described.

The two systems, *A* and *B*, of net subdivision in the ratio 1 : 2 are shown in Figs. 8 and 9, each line between mesh

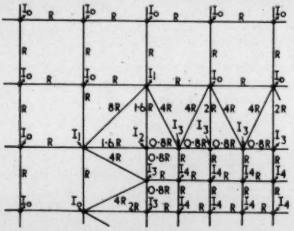


Fig. 8. Subdivision of (x, y)-network in ratio 1 : 2, type *A*

$I_1 = \frac{2}{3}\frac{1}{4} I_0 \qquad I_2 = \frac{1}{3}\frac{1}{2} I_0 \qquad I_3 = \frac{1}{3} I_0 \qquad I_4 = \frac{1}{4} I_0$

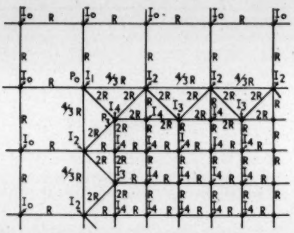


Fig. 9. Subdivision of (x, y)-network in ratio 1 : 2, type *B*

$I_1 = \frac{1}{18} I_0 \qquad I_2 = \frac{1}{9} I_0 \qquad I_3 = \frac{1}{6} I_0 \qquad I_4 = \frac{1}{4} I_0$

points representing a resistance. In each case, diagonal matching resistances are required between the mesh junctions in the coarse net and those in the fine net, but the arrangement is relatively simple, and no intermediate meshes of length $(h/\sqrt{2})$ are required. One of the main advantages is that even along the boundaries of the two nets, including the corner points, there is, as a rule, only a very small loss of accuracy on account of the knitting together of the two different meshes. Mathematical analysis shows that there are certain small error terms which appear along the boundary between the coarse and the fine mesh, but the largest ones of these are always paired in the designs of Figs. 8 and 9. That is, there may be a small error of the kind $+\frac{h^2}{16}\frac{\partial^2 U}{\partial x \partial y}$ at a point at the inner edge of the coarse net, and a compensating error of $-\frac{h^2}{16}\frac{\partial^2 U}{\partial x \partial y}$ at the adjacent point on the outer edge of the fine net. Several of such subdivided (x, y)-networks have been used by the author, and the evaluation of test problems has shown that errors along the boundaries of the coarse and the fine nets due to the matching were of the same magnitude as those due to the resistance tolerances (of the order of one part in 10000).

The design values of the two subdivided networks, Figs. 8 and 9, can be derived either by considering the current flow in the network for certain simple problems, e.g. for the problem of a parallel plate condenser, or from a formal analysis, writing down the Taylor series expansions of the

Fig. 10. Network "star" at point P_0 of Fig. 9

function U at the various network points. Proceeding, in the formal analysis, from the coarse network, one adjusts the weighting values of the various terms, which correspond to the reciprocals of the resistance values, such that the first order terms vanish, and the second order terms represent the Laplacian with the smallest possible error, prescribing "pairing" of errors if these turn out to be unavoidable. This formal method will be illustrated by the evaluation of the resistance values around point P_0 in Fig. 9; the corresponding "network star" is shown in Fig. 10. Using suffixes x and y to denote partial differentiation with respect to x and y, and making use of symmetry conditions, one has:

$$(U_1 - U_0) = -hU_x + \frac{h^2}{2}U_{xx} + 0(h^3)$$

$$\alpha(U_2 - U_0) = -\alpha hU_y + \alpha\frac{h^2}{2}U_{yy} + 0(h^3)$$

$$\beta(U_3 - U_0) = +\frac{\beta}{2}hU_x - \frac{\beta}{2}hU_y + \beta\frac{h^2}{8}U_{xx} - \beta\frac{h^2}{4}U_{xy} + \beta\frac{h^2}{8}U_{yy} + 0(h^3)$$

$$\alpha(U_4 - U_0) = \alpha hU_x + \alpha\frac{h^2}{2}U_{xx} + 0(h^3)$$

$$(U_5 - U_0) = hU_y + \frac{h^2}{2}U_{yy} + 0(h^3)$$

Adding both sides, and choosing the coefficients α and β such that the first order terms vanish, gives the relation $\alpha + \beta/2 = 1$. However, it is not possible to determine the coefficients α and β at the point P_0 separately until similar equations have been found for *all* points along, and adjacent to, the boundary of the two nets; it is then seen that there are in all the same number of independent equations as the total number of coefficients. In the above example, then, one finds $\alpha = \frac{3}{4}$ and $\beta = \frac{1}{2}$ and the right-hand side becomes $\frac{1}{16}\frac{5}{6}h^2\Delta^2 U - (h^2/8)U_{xy} + 0(h^3)$. The error term $-(h^2/8)U_{xy}$ will be found very nearly cancelled by the error term $+(h^2/8)U_{xy}$ arising at the point P_3 (Fig. 9). Hence, the current to be fed in at P_0, in problems governed by equation (1) with g not equal to 0, would be $\frac{1}{16}\frac{5}{6}I_0$, where I_0 is the current $I_0 = -h^2 g_0/R_N$ to be fed in at interior points of the coarse mesh. The described formal method of working out the boundary modifications is tedious, but it gives an insight into the errors which might arise along the join, and it allows the easy determination of the weighting factors of the currents I

to be fed in. In practice, it is found best to combine the two described methods such that the formal method is mainly used to verify the correctness of the weighting coefficients derived from current flow considerations, and the smallness of the error terms. The weighting factors I/I_0 for the currents to be fed in, in Poisson type problems, are the figures given in the legends of Figs. 8 and 9.

The 1 : 2 subdivided network preferred by the author is type *A* (Fig. 8), because a fine net of this type can be easily connected temporarily to an existing network, corresponding to the coarse net, as the values of the resistances along the join which are situated within the coarse net are the standard values for the coarse net. In substituting a subdivided net temporarily, it is therefore only necessary to break the connexions in the coarse net along the boundaries of the area to be subdivided, and to connect the separate fine net with its diagonal resistances already in place, without having to modify at all the resistance values in the coarse net.

Another network subdivision which has been found useful is one in the ratio 1 : 2·5 (or 2 : 5), as its combination with the 1 : 2 subdivision, applied twice, gives a net refinement in the ratio 1 : 10. The required network arrangement is shown in Fig. 11, the values of the resistances and of the

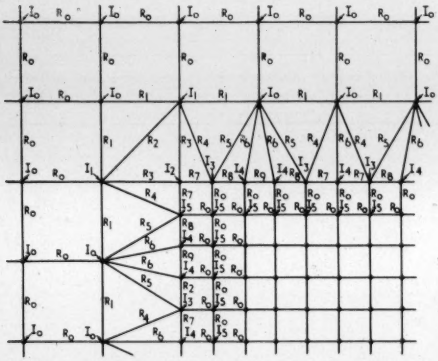


Fig. 11. Subdivision of (x, y)-network in ratio 2 : 5

$R_1 = 1{\cdot}05\ R_0$ $R_2 = 14\ R_0$ $R_3 = 1{\cdot}59\ R_0$ $R_4 = 3{\cdot}33\ R_0$ $R_5 = 10\ R_0$

$R_6 = 2{\cdot}5\ R_0$ $R_7 = 0{\cdot}63\ R_0$ $R_8 = 0{\cdot}58\ R_0$ $R_9 = 0{\cdot}66\ R_0$

$I_1 = 1{\cdot}02\ I_0$ $I_2 = 0{\cdot}44\ I_0$ $I_3 = 0{\cdot}29\ I_0$ $I_4 = 0{\cdot}27\ I_0$ $I_5 = 0{\cdot}16\ I_0$

currents to be fed in at the nodes being given in the legend of the figure. A typical test result, for a small network of 6 × 8 meshes, with an area of 4 × 2 meshes subdivided in the ratio 1 : 2·5, is shown in Fig. 12. The average error expected from the size of the network and the resistance tolerance of 0·2% (see Ref. 3) was about two parts in 10000. It is seen that the average experimental error is even less in this case; the most significant point is that there is no apparent increase in error along the join of the two nets.

Subdivision in the ratio 1 : 2 has also been applied to (r, z)-resistance-networks for solving the Laplace equation, the resistance values required being worked out from the current flow principle, and equation (6). A difficulty arises here at the corner points of the fine mesh, as two different weighting factors are obtained for the diagonal resistance across the corner point, depending on whether one considers it an R_z- or an R_r-resistance. However, unless the off-axis distance of the corner point is a few mesh units only, the difference of the two weighting factors becomes so small

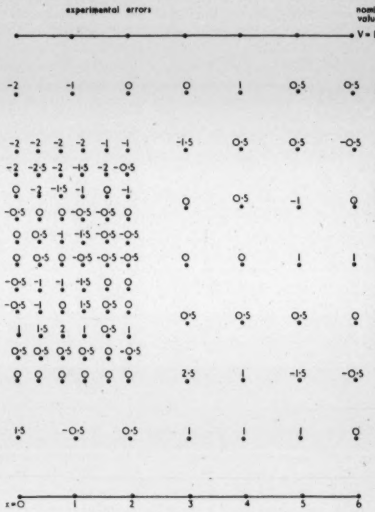


Fig. 12. Typical test result on small (x, y)-network, partly subdivided in ratio 2 : 5, showing that subdivision does not increase errors along join of coarse and fine meshes

that it can be disregarded in practice. In an experimental subdivided (r, z)-network, where the corner point of the fine mesh was only two mesh units distant from the axis, this difference in the required resistance value was found to be 2·8%. Choosing a compromise value led to a subdivided (r, z)-network, in which the errors in evaluated test problems were again found comparable with those due to the resistance tolerances (about one part in 2000 for a coarse network of 6 × 5 meshes).

ACKNOWLEDGEMENTS

The author wishes to thank Mr. A. C. Heath for constructing and testing several types of subdivided networks, and Dr. T. E. Allibone, the Director of the Laboratory, for permission to publish this paper.

REFERENCES

(1) De Packh, D. C. *Rev. Sci. Instrum.*, **18**, p. 798 (1947).
(2) Redshaw, S. C. *Proc. Instn Mech. Engrs*, **159**, p. 25 (1948).
(3) Liebmann, G. *Brit. J. Appl. Phys.*, **1**, p. 92 (1950).
(4) McCann, G. D., and Wilts, C. H. *J. Appl. Mech.*, **16**, p. 247 (1949).
(5) de la Marre, P. H. *C.R. Acad. Sci.* [*Paris*], **235**, p. 602 (1952).
(6) Liebmann, G. *Proc. Instn Elect. Engrs*, **99**, p. 260 (1952).
(7) Allen, D. N. de G., Southwell, R. V., and Vaisey, G. *Proc. Roy. Soc. A*, **183**, p. 258 (1945).

The calibration of circular scales and precision polygons

By A. H. Cook, M.A., Ph.D., National Physical Laboratory, Teddington, Middlesex

[*Paper first received* 20 *January, and in final form* 26 *February*, 1954]

The theory of two methods of calibration is discussed and expressions are obtained for the values and standard deviations of the corrections to the nominal values of the angles. Ways of setting out the computations are also given.

1. INTRODUCTION

The methods used for the calibration of circular scales and precision polygons are very similar in principle to those used for linear scales. (Precision polygons are polygonal prisms with polished faces of hardened steel or glass (aluminized). The angles between the normals to the faces are used to calibrate dividing heads and other instruments in engineering metrology.) Although there is a very extensive literature about linear scales (see Judson[1] for a bibliography) the calibration of circular scales and polygons has had but slight consideration. For this reason and because some of the computations are very much simpler for circular than for linear scales, the theory of the calculations involved in the calibration of a polygon or circular scale is discussed in this paper.

Taylerson[2] has described a method of calibrating precision polygons in which the difference of the angle between two adjacent faces from that between two suitably placed autocollimating telescopes is measured. A similar method was used by Spencer-Jones and Cullen[3] for the calibration of the transit circle of the Royal Observatory, Greenwich. Much greater accuracy is obtained if the angles between all possible combinations of faces are compared, just as in the calibration of linear scales all possible combinations of intervals are compared (for a description see reference to comparators, Johnson[4]). With a circular scale the angles subtended at the centre of the scale by pairs of graduation marks are compared with the angle subtended by two reading microscopes.

2. PRINCIPLE OF METHOD OF COMBINATION OF OBSERVATIONS

Consider a polygon with n faces or a scale with n graduations, labelled 0 to $(n-1)$. Let l_i denote the difference of the angle between the ith and zero face or graduation from the nominal value $2\pi i/n$, and let $\lambda_{(i-j)}$ similarly denote the difference of the auxiliary angle between the autocollimators or microscopes from the nominal value $2\pi(i-j)/n$. The typical observation gives an equation

$$l_i - l_j - \lambda_{(i-j)} = d_{ij} \qquad (1)$$

d_{ij} being the difference of angle obtained from the autocollimator or microscope readings.

From this equation, of which there is one for every possible (integral) value of i and j, the most probable values of the corrections, l, have to be determined. The corrections, λ, to the auxiliary angles are of no direct interest. According to the principle of least squares, these solutions are found by minimizing $\Sigma(l_i - l_j - \lambda_{(i-j)} - d_{ij})^2$ with respect to variations of the l's and λ's. This is essentially the method which was proposed by Hansen[5] for the treatment of linear calibrations, but since the numerical work for linear scales is rather heavy (see Pérard[6]), Thiesen[7] developed a simplified method which is very commonly used (Johnson[4]).

Let α_i be the angle $(l_i - l_{(i-1)})$. On subtracting two successive observation equations, $\lambda_{(i-j)}$ is eliminated:

$$\alpha_i - \alpha_j = d_{ij} - d_{(i-1),(j-1)} = \delta_{ij}, \text{ say.} \qquad (2)$$

The method of solving equations (2) as given by Johnson is equivalent to applying the principle of least squares to these reduced equations, treating them as independent. Now the method of least squares leads to the most probable values of the unknown parameters only if the joint probability of obtaining a set of specified observations is equal to the product of the probabilities of obtaining each of them separately. But if the probability of an observation falling in a given range depends in some way on the values of any of the other observations, then it is no longer necessarily true that the likelihood of obtaining all the observations is a maximum when the sum of squares of residuals is a minimum (see, for example, Jeffreys[8]). Accordingly, the quantities of which the sum of squares is to be minimized must not be correlated if the least squares estimates of the parameters are to be the most probable ones. This condition is not satisfied by the reduced observations, δ_{ij}, for the correlation coefficient of $\delta_{(i+1),(j+1)}$ and δ_{ij} is $\frac{1}{2}$. On the other hand, provided systematic errors have been eliminated, the results of independent observations, d_{ij}, are uncorrelated. The Thiesen method does not, therefore, give the most probable values of the α's, nor in general those of the l's. Pérard[6] has shown that only when the sum of all residuals is zero does the Thiesen solution agree with the most probable solution. A further objection to the Thiesen method is that it does not use all the experimental information. The only reduced equations employed are those obtained from *successive* observation equations (Johnson[4] explains this point clearly), and all other possible reduced equations are neglected.

The great advantage of Thiesen's method in linear calibration is that the calculations are simple and easily remembered. But for circular calibrations, as will be shown, the correct least squares calculations are themselves as simple; and there seems no practical advantage in Thiesen's method to outweigh the theoretical weakness.

3. CALIBRATION OF A SINGLE SCALE OR POLYGON

The observation equations for the calibration of a pentagon are

$$\left.\begin{array}{ll} l_1 - l_0 - \lambda_1 = d_{10} & l_2 - l_0 - \lambda_2 = d_{20} \\ l_2 - l_1 - \lambda_1 = d_{21} & l_3 - l_1 - \lambda_2 = d_{31} \\ l_3 - l_2 - \lambda_1 = d_{32} & l_4 - l_2 - \lambda_2 = d_{42} \\ l_4 - l_3 - \lambda_1 = d_{43} & l_0 - l_3 - \lambda_2 = d_{03} \\ l_0 - l_4 - \lambda_1 = d_{04} & l_1 - l_4 - \lambda_2 = d_{14}. \end{array}\right\} \qquad (3)$$

l_0 is, by definition, zero.

The normal equations are as follows:

For λ_1,

$$5\lambda_1 + \sum_i d_{i,(i-1)} - \sum_i l_i + \sum_i l_{i-1} = 0 \qquad (4)$$

But $\sum_i l_i \equiv \sum_i l_{i-1}$, so that the equation (4) becomes

$$\hat{\lambda}_1 = -\tfrac{1}{3}\sum_i d_{i,(i-1)}. \qquad (5a)$$

Similarly,

$$\hat{\lambda}_2 = -\tfrac{1}{3}\sum_i d_{i(i-2)}. \qquad (5b)$$

The circumflex accent indicates the estimate of the parameter obtained from the actual observations.

For l_1, the normal equation is

$$4l_1 - l_2 - l_3 - l_4 = \Delta_1 \qquad (6a)$$

where $\Delta_1 = d_{10} - d_{21} - d_{31} + d_{14}$

Similarly, the equations for l_2, l_3 and l_4 are

$$-l_1 + 4l_2 - l_3 - l_4 = \Delta_2 \qquad (6b)$$

$$-l_1 - l_2 + 4l_3 - l_4 = \Delta_3 \qquad (6c)$$

$$-l_1 - l_2 - l_3 + 4l_4 = \Delta_4, \qquad (6d)$$

where $\Delta_2 = (d_{21} - d_{32} - d_{42} + d_{20})$, and so on.

A simple formula for the solution of equations (6) is developed below.

The important difference between these equations and those for linear calibration is that the equations for the l's and λ's are independent. This arises because the angle $(2\pi - \theta)$ is numerically equal to θ; one consequence is that the greatest number of autocollimator separations, λ, is $\frac{1}{2}(n-1)$ for n odd or $\frac{1}{2}n$ for n even, instead of $(n-1)$ corresponding parameters for the linear calibration problem.

There is a complication for an even number of faces or graduation marks. When the autocollimators or microscopes are set nominally at π, $\frac{1}{2}n$ observations suffice to include all faces instead of n observations for other settings. But if only $\frac{1}{2}n$ observations are made, the normal equations for the l's and λ's are no longer independent. This is a great disadvantage as it makes the calculations much less simple and it seems worthwhile to avoid it by retaining cyclic symmetry at all intervals and repeating the observations at the interval of π. (Scales are usually read with two microscopes at the ends of a diameter, so eliminating the error due to the centre of rotation of the scale not being the centre of graduation. The observations then repeat after π instead of 2π. In some astronomical work six microscopes are used and the observations repeat after $\pi/3$. This does not affect the theory or practice of the calibration.) Thus for a hexagon the observation equations would be

$$\left.\begin{array}{llll}
l_1 - l_0 - \lambda_1 = d_{10} & l_2 - l_0 - \lambda_2 = d_{20} & l_3 - l_0 - \lambda_3 = d_{30} \\
l_2 - l_1 - \lambda_1 = d_{21} & l_3 - l_1 - \lambda_2 = d_{31} & l_4 - l_1 - \lambda_3 = d_{41} \\
l_3 - l_2 - \lambda^1 = d_{32} & l_4 - l_2 - \lambda_2 = d_{42} & l_5 - l_2 - \lambda_3 = d_{52} \\
l_4 - l_3 - \lambda_1 = d_{43} & l_5 - l_3 - \lambda_2 = d_{53} & l_0 - l_3 - \lambda_3 = d_{03} \\
l_5 - l_4 - \lambda_1 = d_{54} & l_0 - l_4 - \lambda_2 = d_{04} & l_1 - l_4 - \lambda_3 = d_{14} \\
l_0 - l_5 - \lambda_1 = d_{05} & l_1 - l_5 - \lambda_2 = d_{15} & l_2 - l_5 - \lambda_3 = d_{25}
\end{array}\right\} \qquad (7)$$

The normal equations are

$$\lambda_1 = -\tfrac{1}{6}\sum_i d_{i,(i-1)}, \quad \lambda_2 = -\tfrac{1}{6}\sum_i d_{i,(i-2)}, \quad \lambda_3 = -\tfrac{1}{6}\sum_i d_{i,(i-3)} \qquad (8)$$

and

$$\left.\begin{array}{r}
6l_1 - l_2 - l_3 - 2l_4 - l_5 = \Delta_1 \\
-l_1 + 6l_2 - l_3 - l_4 - 2l_5 = \Delta_2 \\
-l_1 - l_2 + 6l_3 - l_4 - l_5 = \Delta_3 \\
-2l_1 - l_2 - l_3 + 6l_4 - l_5 = \Delta_4 \\
-l_1 - 2l_2 - l_3 - l_4 + 6l_5 = \Delta_5,
\end{array}\right\} \qquad (9)$$

where $\Delta_1 = (d_{10} - d_{21} - d_{31} + d_{15} - d_{41} + d_{14})$, etc.

The l equations are slightly more complicated when n is even than when n is odd.

The general form of the normal equations for the l's is

$$\boldsymbol{l}.\boldsymbol{A} = \boldsymbol{\Delta} \quad \text{for } n \text{ odd}$$

and

$$\boldsymbol{l}.\boldsymbol{B} = \boldsymbol{\Delta} \quad \text{for } n \text{ even}.$$

Here $\boldsymbol{l}$ is the row vector $(l_1, l_2 \ldots l_{n-2})$, and $\boldsymbol{\Delta}$ is the column vector $(\Delta_1, \Delta_2 \ldots \Delta_{n-1})$. Consider first the matrix $\boldsymbol{A}$ for n odd. The elements, a_{ij}, are

$$a_{ii} = n, \quad a_{ij} = -1$$

The matrix is of order $(n-1) \times (n-1)$. The elements, A^{ij}, of the reciprocal matrix are easily shown to be

$$A^{ii} = 2/n, \quad A^{ij} = 1/n.$$

The solution, $\hat{l}_i = \Delta_i . A^{ij}$, of the normal equations, is just

$$\hat{l}_i = n^{-1}.\left(\Delta_i + \sum_{j=1}^{n-1}\Delta_j\right); \qquad (10)$$

$$\sum_{j=1}^{n-1}\Delta_j = \sum_{i\,odd} d_{i0} - \sum_{j\,even} d_{0j}. \qquad (11)$$

The variances of the $\hat{l}$'s are proportional to the diagonal elements of $\boldsymbol{A}^{-1}$ and the co-variances to the non-diagonal elements and are therefore:

$$\text{variance of } \hat{l} \qquad : 2\sigma^2/n,$$

$$\text{co-variance of any two } \hat{l}\text{'s} : \sigma^2/n.$$

σ is the standard deviation of a single observation. The number of degrees of freedom used in calculating σ is $\frac{1}{2}(n-1).(n-3)$, that is $\frac{1}{2}n(n-1)$, the number of observations, less $(n-1)$, the number of l's and $\frac{1}{2}(n-1)$ the number of λ's.

The variance of any angle $\hat{\alpha} = \hat{l}_i - \hat{l}_j$ is

$$var\,\hat{\alpha} = 2\,var\,\hat{l} - 2\,cov\,(\hat{l}_i, \hat{l}_j)$$
$$= 2\sigma^2/n. \qquad (12)$$

All intervals on the scale have therefore the same variance.

When n is even the matrix $\boldsymbol{B}$ has the elements

$$b_{ii} = n, \quad b_{i,(i+\frac{1}{2}n)} = b_{(i+\frac{1}{2}n),i} = -2,$$

all other $\quad b_{ij} = -1.$

The elements of the reciprocal matrix are

$$B^{ii} = 2(n+1)/n(n+2), \quad (i \neq \tfrac{1}{2}n)$$

$$B^{\frac{1}{2}n,\frac{1}{2}n} = 2(n+2)^{-1};$$

$$B^{i,\frac{1}{2}n} = B^{\frac{1}{2}n,i} = (n+2)^{-1},$$

$$B^{i,i+\frac{1}{2}n} = B^{i+\frac{1}{2}n,i} = n^{-1},$$

all other $\quad B^{ij} = (n+1)/n(n+2).$

The solutions of the normal equations are therefore

$$\hat{l}_i = \frac{1}{n(n+2)}\cdot[(n+1).\Sigma\Delta + (n+1).\Delta_i - \Delta_{n/2} + \Delta_{i+n/2}]. \qquad i \neq \tfrac{1}{2}n \quad (13a)$$

Here $\quad \sum\Delta \equiv \sum_{all\,i}\Delta_i = \sum_{i=1}^{\frac{1}{2}n} d_{i0} - \sum_{j=\frac{1}{2}n}^{n-1} d_{0j}.$

Also, $\quad \hat{l}_{\frac{1}{2}n} = (\Delta_{\frac{1}{2}n} + \Sigma\Delta)/(n+2). \qquad (13b)$

The variances are

$$var\ \hat{l}_i = 2\sigma^2(n+1)/n(n+2); \quad (i \neq \tfrac{1}{2}n)$$

$$var\ \hat{l}_{\frac{1}{2}n} = 2\sigma^2(n+2)^{-1}.$$

Similarly

$$cov\ (\hat{l}_i, \hat{l}_{\frac{1}{2}n}) = \sigma^2(n+2)^{-1},$$

$$cov\ (\hat{l}_i, \hat{l}_{i+\frac{1}{2}n}) = \sigma^2 . n^{-1},$$

and $$cov\ (\hat{l}_i, \hat{l}_j) = \sigma^2(n+1)./n(n+2). \quad (i,j \neq \tfrac{1}{2}n)$$

Finally, if $\hat{\alpha}_{ij} = \hat{l}_i - \hat{l}_j$,

$$var\ (\hat{\alpha}_{i,\,i+\frac{1}{2}n}) = 2\sigma^2(n+2)^{-1}.$$

and $$var\ (\hat{\alpha}_{ij}) = 2\sigma^2(n+1)/n(n+2). \quad (i,j \neq \tfrac{1}{2}n)$$

For even n, number the of degrees of freedom is $\frac{1}{2}(n-1).(n-2)$.

A convenient way of setting out the computations is as follows.

For each face, except the zero face, form a table:

Measurements on face i

$$\begin{aligned} i - \ \ 0 &= d_{i0} \\ i - \ \ 1 &= d_{i1} \\ &\cdots\cdots \\ i - (n-1) &= d_{i,(n-1)} \end{aligned}$$

Sum, $\Delta_i = d_{i0} + d_{i1} + \ldots + d_{i,(n-1)}$.

Each measurement, d_{ij}, will appear twice in this series of tables, once for face i and once, with opposite sign, for face j. But d_{i0} will appear once only, for face i.

Then for n odd,

$$n\hat{l}_i = \Delta_i + \sum_{all\,j} \Delta_j. \qquad \text{[equation (10)]}$$

and for n even

$$n(n+2).\hat{l}_i = (n+1)\Delta_i + \Delta_{i+\frac{1}{2}n} + (n+1)\sum_{all\,j}\Delta_j - \Delta_{\frac{1}{2}n},$$

except for $i = \frac{1}{2}n$ when

$$(n+2)\hat{l}_{\frac{1}{2}n} = \Delta_{\frac{1}{2}n} + \sum_{all\,j}\Delta_j. \qquad \text{[equations (13)]}$$

No more is needed if the standard deviations are not required.

For a complete solution the auxiliary variables, λ, must be found. Form a second set of tables,

Measurements at interval k

$$\begin{aligned} k \ \ &- \ \ 0 = \\ (k+1) &- \ \ 1 = \\ &\cdots\cdots \\ (k-1) &- (n-1) = \\ &\text{Sum, } n\lambda_k = \end{aligned}$$

Numerical example. The following figures for observations on a hexagon will illustrate the method of calculation. With the notation used above, let the observation equations be

$d_{10} = +1$, $d_{21} = +2$, $d_{32} = -3$, $d_{43} = \ 0$, $d_{54} = +1$, $d_{05} = -1$

$d_{20} = -2$, $d_{31} = +1$, $d_{42} = -1$, $d_{53} = +3$, $d_{04} = -4$, $d_{15} = +2$

$d_{30} = +2$, $d_{41} = -1$, $d_{52} = +3$, $d_{03} = +4$, $d_{14} = -2$, $d_{25} = \ 0$

The units are seconds of arc.

The normal equations for the λ's are

$$\lambda_1 = 0, \quad \lambda_2 = 0{\cdot}17, \quad \lambda_3 = -1{\cdot}0, \quad \text{from equations (8)}$$

Also,

$$\Delta_1 = -1, \Delta_2 = +1, \Delta_3 = -7, \Delta_4 = +3, \Delta_5 = +6, \Sigma\Delta = +2$$

Hence $\hat{l}_1 = 0{\cdot}35$, $\hat{l}_2 = 0{\cdot}71$, $\hat{l}_3 = -0{\cdot}63$, $\hat{l}_4 = 0{\cdot}85$, $\hat{l}_5 = 1{\cdot}33$, from equations (13)

There are ten degrees of freedom, $\sigma^2 = 6{\cdot}35$ and $\sigma = 2{\cdot}52$.

$$var\ \hat{l}_1 = 1{\cdot}91 \qquad var\ \hat{l}_3 = 1{\cdot}59$$

$$\sigma(\hat{l}_1) = 1{\cdot}38 \qquad \sigma(\hat{l}_3) = 1{\cdot}26$$

4. SIMULTANEOUS CALIBRATION OF TWO SCALES OR POLYGONS

Taylerson[2] describes a method of calibrating two polygons simultaneously by comparing the one with the other; circular scales have also been calibrated in this way (Gill[9], Lorentzen[10,11], Jacoby[12]). The method of calculation indicated by Taylerson suffers from the same defects as the Thiesen method discussed above. The least squares treatment is as follows.

Let l_i, m_i denote the differences of the angles between zero and the ith graduation on the two scales from the nominal value $2\pi i/n$ and let μ_k denote the difference of the angle between the zero marks of the two scales from the nominal value $2\pi k/n$. When the zeros of the two scales nominally coincide, the observations of the angles between corresponding graduation marks are

$$l_i - m_i - \mu_0 = d_{ii}.$$

In general, when the angle between the zero marks is nominally $2\pi(i-j)/n$, the observations are

$$l_i - m_j - \mu_{(i-j)} = d_{ij}$$

The normal equations are

$$nl_i - \Sigma m - \Sigma\mu = \sum_{j=0}^{n-1} d_{ij}; \qquad i = 1 \text{ to } (n-1) \quad (14a)$$

$$-\Sigma l + nm_j + \Sigma\mu = -\sum_{i=0}^{n-1} d_{ij}; \qquad j = 1 \text{ to } (n-1) \quad (14b)$$

$$-\Sigma l + \Sigma m + n\mu_k = -\sum_{i} d_{i,(i-k)}; \; k = 0 \text{ to } (n-1) \quad (14c)$$

Here Σl, Σm, $\Sigma\mu$ stand for summations over all l's, m's and μ's respectively. In the last equation k stands for $|i-j|$.

The solutions are

$$\hat{l}_i = n^{-1}.\left(\sum_{j=0} d_{ij} - \sum_{j=0} d_{0j}\right); \quad i = 1 \text{ to } (n-1) \quad (15a)$$

$$\hat{m}_j = -n^{-1}\left(\sum_{i=0} d_{ij} - \sum_{i=0} d_{i0}\right); \quad j = 1 \text{ to } (n-1) \quad (15b)$$

$$\hat{\mu}_k = -n^{-1}\sum_i d_{i,(i-k)} + 2n^{-2}S(d) - n^{-1}\sum_i d_{i0} - n^{-1}\sum_j d_{0j}$$
$$k = 0 \text{ to } (n-1) \quad (15c)$$

$S(d)$ is the sum of all values of d_{ij}, i and j running from 0 to $(n-1)$.

The variances of the l's and m's follow immediately from these solutions.

Thus $$var\ \hat{l}_i = n^{-2}.var\left(\sum_{j=0} d_{ij} - \sum_{j=0} d_{0j}\right) \qquad (16)$$

All the terms in the sums are independent and therefore $var\ \hat{l} = 2\sigma^2.n^{-1}$, where σ is the standard deviation of a single observation, d_{ij}.

Similarly, $cov\ (\hat{l}_i, \hat{l}_j) = \sigma^2.n^{-1}$

Finally, $var\ (\hat{l}_i - \hat{l}_j) = 2\sigma^2.n^{-1}$

The values for the $\hat{m}$'s are the same.

In calculating σ the number of degrees of freedom is $(n-1).(n-2)$, that is, n^2, the number of observations, less $2(n-1)$, the number of l's and m's, and n, the number of μ's.

An arrangement of the computations is as follows.

For each face of each polygon, including the zero faces, form a table:

Measurements on face l_i			*Measurements from face l_0*		
l	m		l	m	
$i -$	0	$= d_{i0}$	$0 -$	0	$= d_{00}$
$i -$	1	$= d_{i1}$	$0 -$	1	$= d_{01}$
. . . . .			. . . . .		
$i - (n-1)$		$= d_{i,(n-1)}$	$0 - (n-1)$		$= d_{0,(n-1)}$
Sum, $\sum_j d_{ij} =$			Sum, $\sum_j d_{0j} =$		

$$n\hat{l}_i = \sum_j d_{ij} - \sum_j d_{0j}$$

Each measurement appears once in an l-table and once, with opposite sign, in an m-table.

If residuals and standard deviations are not wanted, this completes the calculations. To calculate the μ's, form a further set of tables,

Measurements at interval k

l		m	
k	$-$	0	$= d_{k0}$
$(k+1) -$		1	$= d_{(k+1),1}$
. . . . .			
$(k-1) - (n-1)$			$= d_{(k-1),(n-1)}$
		Sum	$= \sum_i d_{i,(i-k)}$

$$S(d) = \sum_k \sum_i d_{i,(i-k)}.$$

$$n\hat{\mu}_k = -\sum d_{i,(i-k)} + 2n^{-1}.S(d) - \sum_j d_{0j} - \sum_i d_{i0}$$

The last two terms are obtained from the previous working.

5. DISCUSSION

The methods of calculation developed above involve no more labour than the Thiesen method and there seems no reason for continuing to use the latter in the two calibration schemes that have been discussed. The first scheme is, however, limited in practice to angles or faces subtending not less than about 20° because of the physical difficulty of setting up microscopes or autocollimators subtending smaller angles at the centre of the scale. The second, comparison, method is not so restricted but the great number of observations involved in a complete calibration at a small interval (324 observations for the eighteen faces or marks at 20° intervals) would usually mean that this scheme also would be applied only to the larger intervals. Just as for linear scales, smaller intervals may be calibrated by subdivision of an interval already examined in the primary calibration. Again, as for linear scales, it will often suffice to calibrate the smaller intervals of just two or three primary intervals. (But a complete calibration is sometimes necessary—Spencer-Jones and Cullen[3].) The method of cross-calibration devised by Guillaume[13] for linear scales, would then be used. Pérard[6] has given the least squares treatment, but so heavy are the computations that, in practice, the Thiesen method of calculation is always used for cross calibration (see Johnson[4]). The differences between the two methods of computation have been examined by Cabrera and Moreau.[14]

Gilet and Watson[15] have published an analysis of linear and circular calibration and have obtained formulae for the variances of estimates, some of which differ from those obtained in this paper. Thus, for the calibration scheme of Section 2 (simple calibration or calibration of the "first type") they obtain the same expression, $2\sigma^2/n$ for the variance of the interval between any two graduations on a circular scale, but their expression for the variance of any single graduation is $r(n-r)\sigma^2/n$, instead of $2\sigma^2/n$ as found in Section 2. This is clearly wrong for no distinction is made between the graduations in the scheme of observations and so, by symmetry, the variance of the estimate of each should be the same and should not vary with the angle from the arbitrary zero. (Except that for an even number of graduations, the π graduation has greater weight than the others in the scheme given in Section 2.) The error in Gilet and Watson's treatment is that they consider the reduced observation equations to be independent, not only in calculating the estimates according to the Thiesen method but also in deriving the variances of the estimates.

They similarly give expressions for the variances of Thiesen estimates of graduation positions for cross-calibration (calibration of the "second type"). These are also in error for the same reason. The correct variances for Thiesen estimates are the following: For the ith graduation mark in a division with n subdivisions: $2\sigma^2(n^3 - n^2 + in - i^2).n^{-4}$. For any interval in the division except the first and last: $2\sigma^2(n^3 - 1).n^{-4}$.

ACKNOWLEDGEMENTS

My attention was called to this problem by Mr. R. S. Marriner and I have had a number of helpful discussions with Mr. J. G. Hayes.

The investigation described was carried out as part of the research programme of the National Physical Laboratory and this paper is published by permission of the Director of the Laboratory.

REFERENCES

(1) Judson, L. V. *Circular No.* 329 (Washington: National Bureau of Standards, 1927).
(2) Taylerson, C. O. *Machine Shop Mag.*, **8**, p. 91 (1947).
(3) Spencer-Jones, H., and Cullen, R. T. *Monthly Notices Roy. Astron. Soc.*, **104**, pp. 218–35 (1944).
(4) Johnson, W. H. *Dict. Appl. Phys.*, **3**, p. 232 (1923).
(5) Hansen, P. A. *Abh. säcks. Ges. (Akad) Wiss. (Math. Phys. Kl.)*, **15**, pp. 524–667 (1873).
(6) Pérard, A. *Trav. Bur. Int. Poids Mes.*, **16**, pp. 1–77 (1917).
(7) Thiesen, M. *Repert. Exp. Phys.* (ed. v. Carl.), **15**, pp. 285–99 and 677–81 (1879).
(8) Jeffreys, H. *Theory of Probability*, 2nd Ed., p. 129 *et seq.* (London: Oxford University Press, 1948).

(9) GILL, SIR D. *Monthly Notices Roy. Astron. Soc.*, **49**, pp. 105–18 (1889).

(10) LORENTZEN, G. *Astron. Nachr.*, **131**, pp. 217–38 (1893).

(11) LORENTZEN, G. *Astron. Nachr.*, **135**, pp. 353–66 (1894).

(12) JACOBY, H. *Astron. Nachr.*, **137**, pp. 357–60 (1895).

(13) GUILLAUME, C. E. *Trav. Bur. Int. Poids Mes.*, **13**, pp. 1–54 (1907).

(14) CABRERA, N., and MOREAU, H. *Trav. Bur. Int. Poids Mes.*, **21**, pp. 1–27 (1951).

(15) GILET, P. M., and WATSON, G. S. *Austral. J. Phys.*, **6**, pp. 155–70 (1953).

The capacity and field of a split cylindrical condenser when the conductors differ in length

By H. J. PEAKE, M.A., M.Sc., and N. DAVY, D.Sc., University of Nottingham

[*Paper received* 1 *February*, 1954]

In this paper $z = b\,\mathrm{dn}(Kw/U)$ is taken as the basic transformation. From this the field and capacity of a split cylindrical condenser are obtained by the method of inversion. By comparing the results of this and a previous paper, two algebraic identities for the elliptic integrals K and K' are obtained.

NOTATION

The co-ordinates of any point in the plane before and after inversion are denoted by $z = x + iy$, and z_1 respectively, $\bar{z}_1$ is the complex conjugate of z_1. The complex potential is $w = u + iv$, where u is the potential and v the stream function. The radius of inversion is "a".

The usual notation is used to denote the Jacobian elliptic functions of modulus k, namely sn$w$, and

$$K = \int_0^{\pi/2} d\phi/(1 - k^2 \sin^2 \phi)^{\frac{1}{2}}$$

and K' is the same function of k' as K is of k.

THE TRANSFORMATION

The transformation*

$$z = b.\mathrm{dn}(Kw/U) \tag{1}$$

gives the complex potential at any point in the z-plane for the case in which a finite plate AB of width $2bk'$, raised to a potential U, is symmetrically placed between two semi-infinite plates CD and EF at zero potential a distance $2b$ apart (Fig. 1). The capacity of this system is $K'/K\pi$.

THE INVERSION

Geometrical argument. On inversion about a point I on the axis of symmetry distant "a" from AB, the arcs A_1B_1, C_1E_1 of the same circle, centre O_1 and radius $a/2$ ($=R$), are obtained.

The arcs OB_1, OE_1 and OP_1 (where P_1 is any point on either arc), subtend angles α, β and θ at O_1 respectively.

It can be shown that

$$b^2/a^2 = \tan^2(\beta/2),$$

$$k'^2b^2/a^2 = \tan^2(\alpha/2),$$

and

$$x^2/a^2 = \tan^2(\theta/2) \tag{2}$$

From these equations we may determine b^2/a^2 and k'^2, for given values of α and β. Indeed,

$$k' = \tan(\alpha/2)\cot(\beta/2) \tag{3}$$

Algebraic transformation. O is taken to be the origin in the z-plane and O_1 that in the z_1-plane. The axis of x is AB and the axis of x_1 is parallel to AB.

By definition of inversion, $IP.IP_1 = a^2$, so that

$$(z + ia).(\bar{z}_1 - ia/2) = a^2 \tag{4}$$

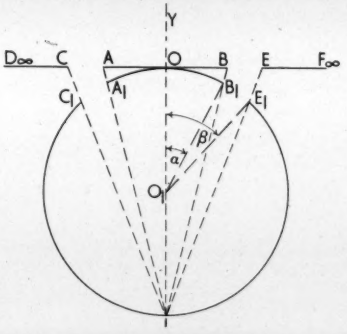


Fig. 1. The inversion

From equations (1) and (4)

$$b\mathrm{dn}(Kw_1/U) = a^2/(\bar{z}_1 - ia/2) - ia$$
$$= -ia(\bar{z}_1 + ia/2)/(\bar{z}_1 - ia/2) \tag{5}$$

If $z_1 = r(\cos\theta + i\sin\theta)$, we obtain after some reduction:

$$b\mathrm{dn}(Kw_1/U) = 2R[2rR\cos\theta + i(R^2 - r^2)]/$$
$$(R^2 + r^2 + 2rR\sin\theta) \tag{6}$$

$$dz/d\bar{z}_1 = -(z + ia)^2/a^2 \tag{7}$$

Electrical (or magnetic), argument. Since inversion is a particular conformal transformation, A_1B_1 and C_1E_1 may be

* THOMSON, J. J. *Recent Researches*, §251, p. 246. (Oxford: University Press, 1893.)

considered to be conductors raised to potential U in the former case and earthed in the latter. Equations (5) and (6) thus give the complex potential w_1 in terms of z_1 at any point in a plane normal to the axis of a split cylinder of radius R. The capacity of the inverted system will be the same as that of the original, namely, $K'/K\pi$.

ELECTRICAL INTENSITY

At any point in the plane the magnitude of the electrical intensity E, is given by

$$E = (1/4\pi)|(dw_1/dz_1)| = (1/4\pi)|(dw/dz)(dz/d\bar{z}_1)|$$

Since $$(d/dw)(\mathrm{dn}w) = -k^2\,\mathrm{sn}w\,\mathrm{cn}w$$

$$dw/dz = -(bKk^2/U)\,\mathrm{sn}(Kw/U)\,\mathrm{cn}(Kw/U)$$

Now, $k^2\mathrm{sn}^2w = 1 - \mathrm{dn}^2w$, and $\mathrm{sn}^2w + \mathrm{cn}^2w = 1$,

so that $$dz/dw = -(K/Ub)\sqrt{[(b^2 - z^2)(z^2 - b^2k'^2)]}$$

Using equation (7), we obtain

$$E = (Ub/4\pi a^2K)|(z + ia)^2|/\sqrt{[(b^2 - z^2)(z^2 - b^2k'^2)]} \quad (8)$$

The magnitude of the electrical intensity is of particular interest on the axes O_1X_1 and O_1Y_1.

For points on O_1Y_1. In this case, $z = iy$, so that $z_1 = R(a - y)i/(a + y)$. Writing t for y_1/R, we have that $y = a(1 - t)/(1 + t)$, and equation (8) becomes

$$E = \frac{(U/2\pi RK)\cos(\alpha/2)\sin(\beta/2)}{\sqrt{[(1 + t^2 - 2t\cos\beta)(1 + t^2 - 2t\cos\alpha)]}} \quad (9)$$

on eliminating b/a and bk'/a by means of relations (2).

For points on O_1X_1. In this case $y_1 = 0$ and if x_1 is replaced by Rt, we obtain for z, $2R(1 - it)/(t - i)$.

After some reduction, we obtain from equation (8), that

$$E = \frac{(U/2\pi RK)\cos(\alpha/2)\sin(\beta/2)}{[(1 + 2t^2\cos 2\alpha + t^4)(1 + 2t^2\cos 2\beta + t^4)]^{\frac{1}{2}}} \quad (10)$$

The field at the centre of the circle has clearly the magnitude

$$(U/2\pi RK)\cos(\alpha/2)\sin(\beta/2) \quad (11)$$

For further details of the algebra the reader is referred to an earlier paper on the same topic.*

SURFACE DENSITY OF CHARGE

At points on a conductor the quantity $(1/4\pi)|dw_1/dz_1|$ is the magnitude of the surface density of charge σ at the point under consideration. Since $z = x$, at such points, we have from equation (8)

$$\sigma = (bU/4\pi a^2K)(a^2 + x^2)/\sqrt{[(x^2 \sim b^2)(x^2 \sim k'^2b^2)]} \quad (12)$$

Using the relations (2), we obtain the expression:

$$\sigma = \frac{(U/4\pi KR)\cos(\alpha/2)\sin(\beta/2)}{\sqrt{[(\cos\theta \sim \cos\alpha)(\cos\theta \sim \cos\beta)]}} \quad (13)$$

CASES OF ESPECIAL INTEREST

Case (i). *One conductor a semi-circle.*

When one of the conductors is a semi-circle: say, $\beta = \pi/2$, we have from equation (3) that $k' = \tan(\alpha/2)$. The above expressions then yield the following results:

Electrical intensity at points on

$$O_1Y_1 \quad \frac{(U/4\pi RK)\sqrt{2}\cos(\alpha/2)}{\sqrt{[(1 + t^2)(1 + t^2 - 2t\cos\alpha)]}} \quad (14)$$

$$O_1X_1 \quad \frac{(U/4\pi RK)\sqrt{2}\cos(\alpha/2)}{[(1 \sim t^2)^2(1 + 2t^2\cos 2\alpha + t^4)]^{\frac{1}{2}}} \quad (15)$$

At the centre, O_1, $E = (U/4\pi RK)\sqrt{2}\cos(\alpha/2)$ (16)

Surface density of charge

$$\sigma = \frac{(U/8\pi RK)\sqrt{2}\cos(\alpha/2)}{\sqrt{[\cos\theta(\cos\theta - \cos\alpha)]}} \quad (17)$$

Tables 1, 2 and 3 give the above quantities for selected values of t and α.

Case (ii). *Symmetrical Case*: *when* $\beta = \pi - \alpha$.

This case has already been investigated in the previous paper. It is of interest to compare the expression deduced from expression (11) above for the magnitude of the field at the centre with that obtained previously. Allowance must be made for the fact that the potential difference was $2U$ then, whereas it is now U.

In the previous case the elliptic modulus k was the quantity defined, namely, $k = \tan^2(\pi/4 - \alpha/2)$; not k' as here.

The comparison yields the identity:

$$\frac{K'[\tan^2(\alpha/2)]}{K[\tan^2(\pi/4 - \alpha/2)]} = \frac{2\sec^2(\pi/4 - \alpha/2)}{\sec^2(\alpha/2)}$$

where $K[\tan^2(\alpha/2)]$ denotes the value of K' for modulus $k = \tan^2(\alpha/2)$.

A similar comparison of the capacity in the two treatments leads to the identity:

K'/K for modulus $\tan^2(\pi/4 - \alpha/2)$ is equal to four times K/K' when the modulus is $\tan^2(\alpha/2)$. Both of these identities may be verified algebraically by means of Landen's transformation.

THE CAPACITY

As has already been mentioned, the capacity of the condenser is $K'/K\pi$. Since $k' = \tan(\alpha/2)\cot(\beta/2)$, we may

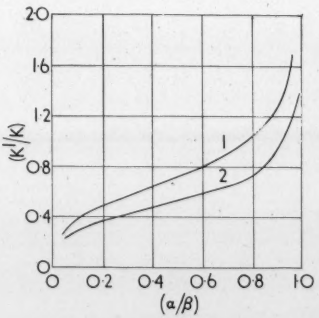


Fig. 2. Graph of (K'/K) against (α/β)
Curve 1, $\beta = 90°$;
Curve 2, $\beta = 150°$.

* PEAKE, H. J., and DAVY, N. *Brit. J. Appl. Phys.*, **5**, p. 316 (1954).

Table 1. *Values of* $(8\pi R/U)E$ *at points on* O_1X_1
where $t = x_1/R$ [using equation (15)]

α \ $\pm t$	0·2	0·4	0·6	0·8	0·9	0·95	1·0	1·05	1·1	1·2	1·4	1·6	1·8	2·0	3·0	5·0
30°	1·01	1·04	1·13	1·39	1·83	2·49	∞	2·31	1·49	1·03	0·63	0·45	0·34	0·27	0·11	0·04
45°	1·15	1·22	1·36	1·72	2·27	3·10	∞	2·87	2·09	1·28	0·77	0·54	0·41	0·32	0·13	0·05
60°	1·24	1·37	1·61	2·15	2·89	3·96	∞	3·66	2·49	1·61	0·95	0·65	0·48	0·37	0·15	0·05
90°	1·33	1·52	1·99	3·54	6·70	12·9	∞	12·3	6·06	2·90	1·33	0·82	0·57	0·43	0·16	0·06

Table 2. *Values of* $(8\pi R/U)E$ *at points on* O_1Y_1'
where, $t = y_1/R$ [using equation (14)]

α \ $\lvert t\rvert$		0·0	0·2	0·4	0·6	0·8	1·0	1·2	1·4	1·6	1·8	2·0	3·0	5·0
Positive t	30°	0·999	1·18	1·36	1·51	1·55	1·37	1·06	0·793	0·596	0·458	0·360	0·144	0·047
	45°	1·12	1·27	1·35	1·35	1·23	1·04	0·834	0·659	0·522	0·419	0·341	0·148	0·051
	60°	1·21	1·29	1·29	1·19	1·03	0·854	0·694	0·562	0·457	0·375	0·319	0·144	0·163
	90°	1·28	1·23	1·10	0·94	0·78	0·64	0·52	0·43	0·36	0·30	0·26	0·13	0·05
Negative t	30°	0·999	0·832	0·681	0·553	0·448	0·366	0·301	0·250	0·210	0·179	0·154	0·081	0·033
	45°	1·12	0·957	0·793	0·648	0·527	0·430	0·353	0·294	0·245	0·209	0·180	0·094	0·038
	60°	1·21	1·06	0·897	0·739	0·603	0·493	0·405	0·336	0·282	0·239	0·204	0·106	0·043

Table 3. *Values of* $(8\pi R/U)E$ *at points on a conductor*
[Using equation (17)]

$(6/\alpha)$	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	0·94	0·98	1·0
						On A_1B_1							
$\alpha =$ 30°	1·37	1·37	1·40	1·44	1·51	1·61	1·76	1·99	2·40	3·35	4·30	7·42	∞
45°	1·04	1·04	1·07	1·11	1·17	1·25	1·39	1·59	1·96	2·79	3·62	6·32	∞
60°	0·854	0·861	0·882	0·921	0·982	1·07	1·21	1·42	1·79	2·72	3·51	6·25	∞
						On either							
90°	0·637	0·645	0·669	0·715	0·787	0·900	1·09	1·40	2·06	4·07	6·77	20·3	∞
						On C_1E_1							
60°	0·493	0·498	0·514	0·542	0·587	0·653	0·755	0·917	1·21	1·88	2·55	4·67	∞
45°	0·430	0·434	0·447	0·470	0·507	0·561	0·643	0·773	1·00	1·53	2·05	3·69	∞
30°	0·366	0·369	0·380	0·399	0·429	0·474	0·540	0·645	0·829	1·25	1·66	2·98	∞

Table 4. *Values of* K'/K

α/β	0·02	0·05	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	0·92	0·94	0·96	0·98	1·0
$\beta =$ 150°	0·248	0·289	0·332	0·391	0·438	0·483	0·531	0·585	0·653	0·750	0·928	0·988	1·07	1·19	1·39	∞
135°	0·260	0·308	0·357	0·425	0·481	0·536	0·592	0·657	0·736	0·847	1·04	1·11	1·20	1·32	1·53	∞
120°	0·271	0·322	0·375	0·451	0·513	0·576	0·639	0·712	0·800	0·921	1·13	1·20	1·29	1·44	1·64	∞
90°	0·296	0·347	0·400	0·486	0·558	0·633	0·707	0·790	0·890	1·03	1·25	1·32	1·41	1·56	1·77	∞

calculate K'/K, either by the method used in the previous paper or by means of tables.

The calculated values of K'/K for selected values of β are given in Table 4. Fig. 2 gives the graphs of K'/K against α/β, when β is 90° and 150°. The graphs for when β is 120° and 135° are similar in shape and lie between the two shown.

It will be observed that these graphs are linear over the range $0{\cdot}2 \leqslant (\alpha/\beta) \leqslant 0{\cdot}5$; i.e. the capacity is directly proportional to the angle α over this range. The following equations give K'/K correct to two places of decimals for the range indicated:

$\beta =$ 90° $(K'/K) = 0{\cdot}342 + (7{\cdot}9)\ 10^{-3}\alpha$ 18° to 45°
120° $= 0{\cdot}324 + (5{\cdot}25)\ 10^{-3}\alpha$ 24° to 60°
135° $= 0{\cdot}313 + (4{\cdot}15)\ 10^{-3}\alpha$ 27° to 67°30′
150° $= 0{\cdot}297 + (3{\cdot}67)\ 10^{-3}\alpha$ 30° to 75°

where α is in degrees.

New books

Twinning and diffusionless transformations in metals. By E. O. HALL, M.Sc., Ph.D. (London: Butterworths Scientific Publications, Ltd., 1954.) Pp. ix + 181. Price 30*s*.

Plastic deformation of metals occurs by slip or by twinning, depending upon the metal and the conditions of deformation. During the last few years much effort has been devoted to the study of slip processes, both experimentally and in terms of dislocation theory. As a result, our understanding of this type of plastic deformation has advanced considerably. By comparison, the process of twinning has been relatively neglected, but the time is now ripe for a more concerted attack upon this problem and upon the related one of shear transformations. It is with this in mind that the present book has been written. Some forty pages are devoted to a preliminary consideration of crystallography and deformation and are intended to prepare the reader for the later parts of the book. Those who need such preparation are hardly likely to be seriously interested in twinning; those who do not may regard the first two chapters as an unnecessary luxury in these days of high book costs. It is perhaps ungracious to make this criticism, for Dr. Hall has reviewed the published work on twinning and related phenomena in an admirable fashion. His arrangement is systematic and the clearly written text is supported by numerous well-chosen illustrations. The book will be welcomed by those looking for a competent and interesting account of present knowledge in this field.

A. G. QUARRELL

Industrial electronics. By R. KRETZMANN. Translated by H. Carter. (London: Cleaver Hulme Press Ltd., 1953.) Pp. 235. Price 25*s*.

Intended primarily as a guide to the application of electronics to mechanical and electrical engineering, this book will also be of wider interest in view of the greater emphasis on thyratrons and other gas-filled devices than is usual in introductory books of electronics.

The first third of the book is devoted to electronic tubes and their elementary associated circuits. The uses of semiconductor devices such as transistors and crystal diodes are not discussed. The remainder of the book deals with the application of electronic tubes to particular techniques in mechanical and electrical engineering. These applications are included under such headings as electronic relays and timers, industrial rectifier circuits, electronic dimming, speed and temperature control, electronic control of resistance welding and electronic motor control. There are also chapters on inductive and dielectric heating.

In a book of this nature, published by an industrial firm, there is a tendency for the viewpoint to become narrow. Such devices as magnetic amplifiers and servomechanisms are among those omitted, and together with the absence of a comprehensive bibliography the result is a rather badly balanced book. The reader will therefore have to look elsewhere for a general introduction to the application of electronics to industry. Those devices which are described are adequately explained and illustrated and assume only a moderate technical knowledge by the reader.

The book is well presented and has few errors, the translation is likewise good apart from the occasional unfamiliar construction.

R. C. HART

Methods of algebraic geometry. Vol. 3. Birational geometry. By W. V. D. HODGE, Sc.D., F.R.S., and D. PEDOE, B.Sc. (London: Cambridge University Press, 1954.) Pp. x + 336. Price 40*s*.

This is the third of a series of books, with the same general title by the same authors, the senior of whom is Lowndean Professor at Cambridge. The two earlier volumes were published in 1947 and 1952. All three volumes are concerned with the properties of one or more sets of numbers $(x_0, x_1, \ldots x_n)$ which satisfy a group of simultaneous equations

$$f_i(x_0, x_1, \ldots x_n) = 0 \qquad (i = 1, 2, \ldots k)$$

Such numbers may be thought of as co-ordinates of a point lying on the intersection of k generalized surfaces $f_i = 0$. They are referred to as an algebraic variety. The present volume is concerned more particularly with two such varieties x and y, between which there exist various equations of the form

$$y_j f_j(x_0, \ldots x_n) - y_i f_j(x_0 \ldots x_n) = 0$$
$$x_j g_i(y_0 \ldots y_m) - x_i g_j(y_0 \ldots y_m) = 0$$

The varieties are then said to be in birational correspondence. The properties of such birational correspondences are investigated in some detail, with particular reference to the ideal theory of rings and the theory of valuation. The analysis, naturally enough, is purely formal, though the style and format are straightforward and clear. It is not easy to see just where this kind of work is useful in theoretical physics. But without doubt there must be some uses, if only we know what they are.

C. A. COULSON

Temperature measurement in engineering, Vol. 1. By H. DEAN BAKER, Ph.D., E. A. RYDER, M.E., and N. H. BAKER, M.A. (London: Chapman and Hall Ltd.; New York: John Wiley and Sons Inc., 1954.) Pp. vi + 179. Price 30*s*.

This book is the first of two volumes, and, after an introduction dealing with the temperature scale, is confined to the industrial use of thermocouples.

A chapter on heat transfer calculations for use in the design of thermocouple sheaths is a notable feature for, as the authors point out, the engineering practice of design "based on quantitative calculation from established facts, as supplemented by judgment" has, in the field of temperature measurement, "all too frequently been curtailed to the exercise of judgment as tempered by optimism." Provided that the need for the use of judgment is not overlooked, the various graphs given in this section should be helpful.

The book contains a great deal of useful information of a practical kind, but the treatment as a whole is curiously unbalanced. For example, twenty pages are devoted to the drilling of small holes, while the whole question of indicators, potentiometers and recorders is dismissed in three. However, there is a large bibliography (mainly, but not entirely, of American references) and its use should enable the reader to fill in most of the gaps.

One error should be pointed out: the Ninth General Conference of Weights and Measures did not recommend a value of 273·15° K for the ice point on the Kelvin scale. There is, at present, no internationally agreed value.

J. A. HALL

Advances in electronics, Vol. 5. Edited by L. MARTON. (New York: Academic Press Inc., 1953.) Pp. xi + 420. Price $9.50.

There are eight review articles in this present volume so that only brief mention can be made of each in a short notice. This volume contains also a cumulative index of main subject titles and authors for Volumes 1 to 5.

In "Performance of visible and infra-red detectors" R. C. Jones compares the behaviour of the several devices at signals close to their noise levels; this is a timely review with 203 references. "Beta-ray spectrometers," by R. W. Hayward is a description of the fundamentals and characteristics of certain instruments of both prismatic and lens types.

Three years ago, in this series, Garlick reviewed the field of solid-state luminescence. Much interesting and practical activity has been taking place in this subject and the review is brought up to date by F. E. Williams. In an attempt to provide, in certain applications, more rugged emitting cathodes than are possible with the usual barium strontium oxides, thorium oxide is being experimented with and details are given by W. E. Danforth.

H. C. Weingartner and S. W. Kennedy provide "A review of modern vacuum pumps in electronic manufacturing." For the sake of completeness one would, perhaps, have liked to see the review cover the complete pumping system rather than the pumps alone. In "A review of recent work in color television" G. J. Hirsch describes the position in the U.S.A. as at June 1953 and deals, mainly, with the "compatible" system.

Though the article by J. S. Schaffner is entitled "Junction transistor applications," it is more concerned with basic circuits for these devices and these, generally, for low-frequency applications in accordance with the present state of the art. The remaining article by R. Q. Twiss "On the steady-state theory of the magnetron" is more in the nature of an original paper by the author and might well have been presented to a learned society; it deals with conditions when account is taken of the emission velocities of the electrons from the cathode.

A. J. MADDOCK

The printing of mathematics. By T. W. CHAUNDY, P. R. BARRETT and C. BATEY. (London: Geoffrey Cumberlege; Oxford University Press, 1954.) Pp. x + 105. Price 15*s*.

This is an important book. To follow its recommendations, born of experience, will not only ease the task of authors, editors, and printers, but will make the final result more readily understandable by the reader, and may well save expense—a need felt by most publishers these days. Its authors are a Reader in Mathematics at Oxford (Chaundy), the Mathematical Reader at the University Press (Barrett) and the Printer to the University (Batey). The improvement in mathematical printing in the last few years owes a good deal to them and their colleagues in other university presses, and to the Monotype Corporation.

The first chapter is a brief introduction to the mechanics of mathematical printing, particularly by the Monotype machine. The second chapter, which forms the main part of the book, is divided into sixteen sections giving detailed recommendations. While I think editors would accept the majority of these—"the accepted framework of mathematics"—other suggestions would be questioned. But, as the authors remark about what they have written, ". . . some is frankly new and experimental; some even provocative; sometimes decisions have been made merely to secure uniformity; . . . there may be conflict with what is recommended in other guides." The use of mathematics in scientific publications has increased rapidly in recent years and so has the number of special symbols required. As the authors write, "The tradition of mathematical printing is still in the making and it would be a pity to clamp down too soon on every convention."

The third chapter contains the technical rules for the Oxford compositors of mathematics; they have been evolved over the last twenty years. There are three appendixes: legible handwriting; type specimens and list of special sorts of type; abbreviations.

This book should be in every college and laboratory library in which prospective authors may be found, so that it is readily available to them to peruse before they start writing mathematics for printing, and to use as a reference book. If they make a habit of writing mathematics for publication they should have their own copies on their desks; and so should editors. Perhaps it is not too much to expect that compositors and readers in other presses will also refer frequently to this book, but without necessarily accepting all its detailed recommendations.

H. R. LANG

Millimicrosecond pulse techniques. By I. A. D. LEWIS, M.A., A.Inst.P., Grad.I.E.E., and F. H. WELLS, M.Sc.Eng., A.M.I.E.E. (London: Pergamon Press Ltd., 1954.) Pp. xiv + 310. Price 40*s*.

The development of pulse techniques in the millimicrosecond region is of growing interest in some branches of physics and electronic engineering. This book, therefore, makes a timely appearance in a field which is undoubtedly of a specialized nature and consequently not familiar to many. Nevertheless, it is becoming increasingly necessary for physicists and electronic engineers to acquaint themselves with the advances in millimicrosecond pulse techniques as they become aware of the limitations imposed by conventional circuits. The subject-matter of this book will be of use primarily to those already familiar with pulse techniques, but it will also have an appeal to the general physicist and electronic engineer. An excellent balance between theory and practice has been achieved.

The book opens with a short theoretical introduction on circuit theory and some advanced mathematics employed in circuit analysis. It then deals with the theory of transmission lines and consideration is given to helical and lumped delay lines which have special applications in the millimicrosecond region. A chapter is devoted to line transformers which are dealt with in some detail. Unfortunately, the treatment of pulse transformers is very superficial, though the reader is referred to a set of references at the end of the book!

Various circuits employed in pulse generators, amplifiers and cathode-ray oscillography are treated admirably in three following chapters. The limitations of existing circuits are pointed out, together with the new techniques employed to overcome them. These techniques include the use of secondary emission valves, distributed amplification and cathode-ray tubes employing transmission line deflexion systems. Descriptions of various practical circuits are given and they provide a ready source of information to the practical designer.

The book concludes with two chapters on the applications of millimicrosecond pulse techniques, one of which is devoted to the field of nuclear physics in which the authors have a special interest. Finally, the bibliography and a numerous set of references at the end of the book provide a valuable source of more detailed information on specific subjects.

F. R. CONNOR

Notes and comments

Notes for authors

The Institute of Physics has just published a revised edition of its 32-page pamphlet *Notes on the preparation of contributions to the Institute's Journals and other publications*. It is intended to assist less experienced authors and to serve as a reference booklet for all who wish to contribute to the Institute's publications. The new edition gives hints on the preparation of a script and diagrams, on the lay-out of mathematics for the printer, the correction of proofs and so on. In addition to a short bibliography of reference books and works on technical writing, there are lists of the spellings, symbols and abbreviations used by the Institute. These have been modified a little in this new edition to conform with the new British Standard 1991, Part 1: 1954, *Letter symbols, signs and abbreviations, Part* 1, *General*, which authors of mathematical papers are also advised to consult.

Copies of the revised pamphlet may be obtained from The Institute of Physics for 2*s*. 6*d*. (including postage).

Communications and Electronics

We have received the first issue of a new monthly trade journal entitled *Communications and Electronics*, the aim of which is stated to be "to provide a link between the designer and manufacturer of British communications and electronic equipment on the one hand and the user on the other." A special appeal is to be made to users of British communications and electronic equipment overseas. The emphasis throughout is on the practical applications of new communications and electronic techniques, rather than on theory and circuit design. The issue before us is well and pleasingly produced and printed, and its 90 pages of text, measuring $11\frac{1}{2} \times 8\frac{3}{4}$ in., are copiously illustrated. We wish our new contemporary success. The publishers are Heywood and Co. Ltd., one of the National Trade Press Group, and the subscription rate is £1 10*s*. per annum, $6.00 in U.S.A. and Canada.

Welding polythene and polyvinyl chloride

A nine-page booklet describes the welding of polythene and polyvinyl chloride using nitrogen in a hot gas process. The information in the booklet is based on an extract from *Chemical and process engineering*, published by Leonard Hill Ltd., London.

Copies of the booklet may be obtained free of charge from The Edison Swan Electric Co. Ltd., 155 Charing Cross Road, London, W.C.2.

Radiation hazards in industry

The British Occupational Hygiene Society is holding a conference at the London School of Hygiene and Tropical Medicine on Monday, 1 November, 1954. The conference will be devoted to a discussion of radiation hazards in industry. Forms of application may be obtained from the society's Honorary Secretary, Mr. P. C. G. Isaac, Public Health Engineering Laboratory, King's College, Newcastle upon Tyne.

ERRATUM

In the paper *Densitometric evaluation of coal-dust stains on filter paper* by J. G. Dawes, which was published in the June issue of this *Journal* (pp. 221–224), the scale under Fig. 3 corresponds to 0 to 100 μ, not to 0 to 5 μ.

British Journal of Applied Physics

Original contributions accepted for publication in future issues of this Journal

Further tests on the stability of analytical weights in chemical laboratories. By P. H. Bigg and F. H. Burch.
The two-dimensional magnetic or electric field of a single isolated pole-piece. By N. H. Langton and N. Davy.
The field due to an infinite dielectric cylinder between two parallel conducting planes. By C. Mack.
Heat transfer through oxide cathode materials. By A. E. Pengelly.
A note on the theory of the rubber membrane. By W. Fulop.
The calculation of heat and flow through disks and its application to conductivity measurements. By J. C. Jaeger and A. Beck.
The use of semiconductors in thermo-electric refrigerators. By H. J. Goldsmid and R. W. Douglas.
The mechanical impedance of the human mastoid process. By R. S. Dadson, D. W. Robinson and R. G. F. Greig.
Comparison of radioactivities by the use of X-ray film. By A. N. Davenport and G. W. W. Stevens.
Growth of barium orthosilicate interface of oxide-coated cathodes. By M. G. Harwood and Nora Fry.
Determination of thermal conductivities at high temperatures. By K. S. Krishnan and S. C. Jain.
The resistance of an elliptic plate. By E. E. Jones.
Ionic bombardment heating of magnetron cathodes. By A. E. Barrington.
Thulium 170 for industrial radiography. By R. Halmshaw.
The calculation of voltage surges in a Van de Graaff generator. By B. Millar.
A method of making silicon junction diodes. By J. W. Granville.
Elastic work involved in rolling a sphere on another surface. By D. Tabor.
Some lithium iodide phosphors for slow neutron detection. By K. P. Nicholson and G. F. Snelling.

Journal of Scientific Instruments

Contents of the October issue

SPECIAL ARTICLE

The physics of cathode-ray storage tubes. By C. N. W. Litting.

ORIGINAL CONTRIBUTIONS

Papers

A phonic wheel generator for position indication in digital computor magnetic drum storage. By D. R. Quested and A. D. Booth.
A nephelometer of wide range for bacteriological use. By E. O. Powell.
The radio-frequency system of the Birmingham proton synchrotron. By L. U. Hibbard.
An instrument for measuring energies and angular distributions of charged reaction products. By C. H. Westcott, H. I. S. Allwood, J. N. Dodd, D. H. Simmons and C. J. Baker.
A filter photometer for use at sea. By R. A. Cox.
Statistics of photomultiplier scintillation counters. By G. T. Wright.
The use of double triodes in high-gain, low-frequency amplifiers. By R. G. Wylie.
A combined dilatometer and electrical resistivity apparatus for studies in powder metallurgy. By N. A. McKinnon.

Laboratory and workshop notes

A device for maintaining a small pressure against a steady leak. By J. H. Prentice.
A laboratory pump for closed circulation of humidified air. By C. Wood.
A shunt-compensated stabilizer for microdensitometer light sources. By E. Cohen.
Metal to glass seals for vacuum work at low temperatures. By J. E. Quarrington.
A gas leak. By J. Flinta.
Aids to X-ray fibre-photography. By L. Brown and A. W. Porter.

NOTES AND COMMENTS

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone*: Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions*. (Price 2*s*. 6*d*. including postage.)
REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone*: Monarch 7644.)
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.
SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s*. each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

The near infra-red absorption spectra of natural and synthetic fibres*

By A. Elliott, D.Sc., W. E. Hanby, B.Sc., A.R.I.C., and B. R. Malcolm, Ph.D., Courtaulds, Ltd., Maidenhead, Berks.

The near infra-red region may be used to obtain spectra with polarized radiation of natural and synthetic fibres. This region has considerable experimental advantages and the spectra of poly-*L*-alanine, polyglycine, wool, silks and commercial regenerated protein fibres are shown and discussed. Besides supplementing other types of observation, information may be obtained on the orientation and molecular configuration not always given by other methods.

The infra-red absorption spectrum is capable of providing considerable information regarding the composition, degree of molecular orientation and in some cases amount of crystallinity of fibres. In the range 3 to 10 μ wavelengths, the thickness of the specimen necessary to obtain the proper strength of absorption is restricted; in the case of protein fibres, for example, the specimen should usually be not more than 5 μ thick. Thicker specimens have to be reduced either with a microtome or by grinding. The width of the specimen is also important. In order that the image of the specimen shall fill a 0·1 mm wide spectrometer slit a circular fibre 5 μ in diameter has to be magnified at least 20 times and in actual practice considerably higher magnification is desirable. The design of a suitable instrument presents considerable problems. Moreover, in order to fill the aperture of the spectrometer it is necessary to use a wide cone of radiation from the specimen. When used with polarized radiation the effect of this is to lower the dichroism and corrections are difficult to apply. This is therefore a region which is not at all suitable for the rapid examination of fibres and raises considerable experimental difficulties.

In the region 1–2·5 μ the bands are either overtones or combinations of the fundamental vibrations, with much lower absorption coefficients so that a specimen may be about 0·2 mm in diameter. These may be used with a reflecting microscope of about ×5 magnification. Thinner specimens can be examined by tying a number of fibres together in a small bundle and immersing them in a suitable transparent liquid of about the same refractive index, to reduce the amount of scattered radiation.

From an experimental standpoint this region has other considerable advantages. The source may be a tungsten filament lamp in a glass envelope. The spectrometer prism can be of second quality fused quartz which has a high dispersion and is quite inexpensive. A lead sulphide cell may be used as the detector with an 800 c/s amplifier. This avoids many of the troubles commonly associated with low-frequency amplifiers for thermocouples. Cells for liquids can be made from glass, and since the wavelength range is fairly close to the energy peak of the source, trouble from stray radiation is seldom serious. These advantages are considerable and make the exploration of this region for the examination of fibres well worth while.

The most serious disadvantage of this region is that at our present state of knowledge it is hardly ever possible to be certain of the origin of all the bands, which are often superposed one on another so that quantitative measurements on them are less reliable. An empirical approach can, however, yield considerable information in certain cases and the difficulties mentioned are not entirely avoided when working in the fundamental region.

The type of information which can be derived from a study of the spectra in the 1–2·5 μ range is best illustrated by the examples which follow. Most of these refer to protein fibres or synthetic polypeptides. While these appear particularly interesting there is no reason *a priori* to assume that a similar wide investigation of other types of fibre might not also provide information of equal interest.

One of the most useful applications of this region is measurements on the dichroism of bands in order to estimate the degree of molecular orientation in the fibre. In this respect the measurement is usually taken over both the crystalline and amorphous regions since normally these are undifferentiated in the spectra. Infra-red observations therefore are complementary to an X-ray examination which reveals the orientation only of the crystalline regions and the two may be profitably combined.

EXPERIMENTAL METHODS

Many of the features of apparatus used to obtain the results which follow have already been mentioned. We have used either a Perkin–Elmer single-beam spectrometer with a lithium fluoride prism or a spectrometer constructed by ourselves with a fused quartz prism. This instrument has two spherical mirrors arranged in the manner shown by Tetlow, McAuslan, Brinley and Price[1] after the method recommended by Czerny and Turner[2] and avoids the use of an off-axis parabolic mirror. Both instruments are fitted with reflecting microscopes giving ×5 magnification and the specimen is moved in and out of the beam at two-second intervals. In this way the spectrum of the specimen and the incident beam are obtained almost simultaneously and errors caused by long period fluctuations in the apparatus are minimized. This method of recording contributes greatly to the accuracy of the result and is of great assistance in observing small differences in otherwise similar spectra such as are produced when the dichroism is low. The polarizer used has been described by Elliott, Ambrose and Temple[3,4] and the recorder is a modification of that described by Elliott and Ambrose.[5]

When measurements are being made on a large number of fibres which are optically inhomogeneous it is often impossible to reduce the scatter by means of an immersion liquid to a sufficiently small amount for it to be comparable with the true absorption. In such cases the reference beam may be defocused by placing a piece of glass over the part of the cell through which the reference beam passes. If I is the

* This article was read at a meeting of the Industrial Spectroscopy Group of The Institute of Physics on 29 April, 1954.

transmitted intensity, I_0 the incident intensity, a and s the absorption and scattering coefficients, then

$$I = I_0 \exp\left[-(a+s)d\right]$$

where d is the thickness of the specimen. In general both a and s depend on the direction of the electric vector relative to the fibre axis. If the reference beam is reduced in intensity by a fraction k we obtain

$$\log(kI_0/I) = (a+s)d$$

$\text{Log}_{10}(I_0/I)$ is the optical density and is measured from the spectrometer record. Since both s and k vary only slowly with wavelength compared with a, the absorption bands appear superposed on an arbitrary, even background. This may be estimated if there is a region in which the fibre is known to be transparent and subtracted if necessary. For many purposes, however, all that is required is to compare the shape and relative intensities of bands and in this case the absolute zero does not need to be estimated.

The dichroic ratio of a band in a fibre is measured by the ratio of the absorption coefficients parallel and perpendicular to the fibre axis. It depends on the angle α of the transition moment giving rise to the absorption, relative to the molecular axis and the orientation of the molecules relative to the fibre axis. The dependence of the dichroic ratio on α and on the average angle of orientation has been calculated by Elliott, Ambrose and Temple[6] for $\alpha = 0°$ and $90°$. Qualitative comparisons between fibres of the same type can be made even if α is unknown since dichroism cannot arise without molecular orientation. It does not follow, however, that since a fibre is not dichroic that there is not a regular molecular structure, as for example in some biological structures where layers of fibres may develop in a criss-cross pattern.

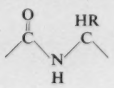
GENERAL FEATURES OF OVERTONE SPECTRA

Most fibres contain some water which gives rise to a strong band at about 5150 cm^{-1}. Quite small amounts are detectable so that the water uptake of textile fibres can be measured quantitatively.[7] Examination of the infra-red spectrum is probably the most reliable way of finding out whether a fibre has any residual water in it after drying.

In the range 4350–4450 cm^{-1} there are often quite sharp bands which are combinations of CH stretching and deformation modes. At lower wave numbers, 4350–3900 cm^{-1}, while often there are a number of resolvable peaks, their assignment is difficult since they appear to arise from combinations with skeletal modes.

Fibrous proteins of the keratin type have long been known to exist in two forms; in the α-form the polypeptide chain is folded or coiled in some way whereas in the β-form the chains are much more nearly extended (Astbury and Street[9]). More recently, synthetic polypeptides have been shown to occur in two forms which have many of the characteristics of α- and β-keratin (Bamford, Hanby and Happey[10]; Ambrose and Elliott[11]). These forms are also referred to as α and β, though it is recognized that the polypeptide chain configurations are not identical with the corresponding forms of the fibrous proteins.

The repeating unit of the polypeptide chain is



where R represents a side chain. In the synthetic polypeptides polyalanine and polyglycine R is CH_3 and H respectively and natural proteins are characterized by a wide variety of side-chains. Silks are particularly simple proteins consisting mainly of alanine and glycine. It has been shown by Ambrose and Elliott[12] that the two forms α and β may be distinguished by two absorption bands of the polypeptide chain in the overtone region. The first of these is a band at 4850 cm^{-1} which has high parallel dichroism in extended (β) polypeptides and perpendicular in folded (α) polypeptide chains. This is probably the same band as that observed in nylon by Glatt and Ellis[13] at 4883 cm^{-1} and is assigned by them to a combination of NH stretching and in-plane deformation modes. The second band is at about 4520 cm^{-1} in the β-form and at about 4600 cm^{-1} in α polypeptides and proteins. The band at 4520 cm^{-1} shows parallel dichroism comparable with the NH combination, but dichroism is usually not observed in oriented α-fibres in the 4600 cm^{-1} band. Reasons have been given for assigning this band to a CO combination mode[12] but the fundamental frequencies have never been identified. Recent experiments on the deuteration of synthetic polypeptides have shown that on replacing the amide hydrogen by deuterium this band disappears at the same rate as the NH combination band. It therefore seems unlikely that it is a simple CO combination mode and probably directly involves the NH group. Price and Fraser[14] have suggested that a band at 1270 cm^{-1} in the spectrum of α-proteins is predominantly a C–N vibration and it seems possible that this couples with the NH stretching band at 3300 cm^{-1} to give the band at 4600 cm^{-1}. This is supported by the disappearance of this band and the development of a band at 1220 cm^{-1} in poly-*L*-alanine when it undergoes the α- to β-transformation, which would be consistent with the observed frequency shift of the combination. This will be discussed more fully elsewhere; whatever the precise origin of this band, its empirical correlation with the α- and β-configurations is well established and is consistent with infra-red observations on the fundamental CO and NH modes and X-ray diffraction results. These results have been applied by Elliott to an examination of the stretching of hair[15] and by Ambrose and Elliott [16,17] to globular proteins.

There is a remarkable agreement between the main frequencies of the globular proteins and the synthetic α-polypeptides and the frequency shifts on denaturation corresponding to a partial transition to the β-form. This cannot, however, be taken to imply that the α-configuration is necessarily the same in globular proteins as in synthetic polypeptides with inert side chains, where a helical structure is consistent with most of the experimental results. As yet no crystalline globular protein has been found to give high infra-red dichroism and X-ray studies have not so far given an unambiguous result. This must be remembered when comparing the spectra of synthetic and polypeptide fibres with regenerated protein fibres.

EXAMPLES OF INFRA-RED SPECTRA OF FIBRES

Poly-L-alanine.

α-form: Fibres of poly-*L*-alanine which contain a small amount of dichloracetic acid may be oriented by stretching, and are then found to be predominantly in the α-form, though a small amount of the oriented and crystalline β-form is also produced. The fibres may be suitably conditioned by soaking in a solution containing 12 parts of dichloracetic acid in 100 parts of carbon tetrachloride (by volume), and then

drying in air for about 15 min in order to allow the carbon tetrachloride to evaporate. Cold drawing is accompanied by very pronounced "necking," and it is noteworthy that even quite irregular fibres may be successfully drawn in this way. The extension possible is fairly constant, about 170%. The dichloracetic acid may be removed by heating or by washing, for example in ether or carbon tetrachloride. Removal by heat produces a very crystalline specimen, from which a very well-ordered X-ray diagram may be obtained [see Bamford, Brown, Elliott, Hanby and Trotter[18] Fig. 1(*a*)]. Washing out the acid produces a much less crystalline fibre, but the infra-red spectrum appears to be the same in both cases, except for the fact that the water band at approximately 5250 cm^{-1} is much reduced in the spectrum of the heated specimen.

The spectrum of a heated specimen is shown in Fig. 1.

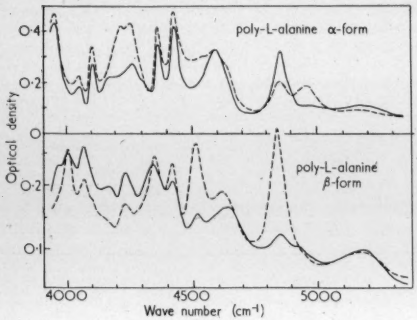


Fig. 1. Spectra of oriented fibres of α (folded) and β (extended) poly-*L*-alanine observed with polarized infra-red radiation

Full line—E vector perpendicular to fibre axis.
Broken line—E vector parallel to fibre axis.

β-form: Cold drawing of an air dried, poly-*L*-alanine fibre results in the production of an oriented β-fraction, with a roughly equal amount of slightly oriented α-material. Similar results are obtained if a water-soaked fibre is drawn at room temperature. If, however, the stretching is carried out in steam, a greater proportion of the polymer is extended into the β-form and high crystallinity and orientation of this form is obtained [Bamford and others[18] Fig. 1(*b*)]. The spectrum of a poly-*L*-alanine fibre stretched in steam is shown in Fig. 1.

The fact that the predominance of (respectively) α- and β-forms of the polymer has been convincingly demonstrated by the methods of X-ray diffraction greatly strengthens the validity of the interpretation of the spectra shown in Fig. 1. In particular, the association of the parallel band at 4510 cm^{-1} with a β-configuration and of the non-dichroic band at 4600 cm^{-1} with the α-form are confirmed. In polyalanine, there is also a small frequency change in the combination band from 4845 cm^{-1} (β) to 4860 cm^{-1} (α), though this shift is too small to be of value for diagnostic purposes.

The bands at 4365 cm^{-1} and 4425 cm^{-1} in α-poly-*L*-alanine, and corresponding bands at a slightly lower wave number in the β-form are almost certainly CH_3 combinations of stretching and deformation modes, as may be seen by comparing them with bands in the spectrum of methylene chloride.[19] Some of the bands at lower wave numbers show quite high dichroism, but as yet nothing is known of their origin. The very highly dichroic band at 4965 cm^{-1}, noted by Glatt and Ellis[20] as appearing in some nylon specimens, is very prominent in the spectrum of α-poly-*L*-alanine, but hardly to be seen in the β-form. Like the band at approximately 4850 cm^{-1}, its dichroism changes with change of configuration from α to β. If, as seems likely, the 4965 cm^{-1} band is associated with crystal lattice modes, it is nevertheless not possible to correlate its strength with the "crystallinity" of a polymer as deduced from the X-ray diffraction diagram. In support of this statement, it may be mentioned that oriented α-poly-*L*-alanine which has not been heated (see above) shows the band quite strongly: it is also prominent in poly-*DL*-leucine.[8] However, the appearance of an X-ray diagram depends to a great extent on crystallite size (and, of course, orientation) and possibly infra-red bands would be produced by much smaller crystal regions than would be needed to produce sharp X-ray reflexions.

Polyglycine.

The specimen of polyglycine (Fig. 2) was not of sufficiently high molecular weight to enable it to be oriented and was

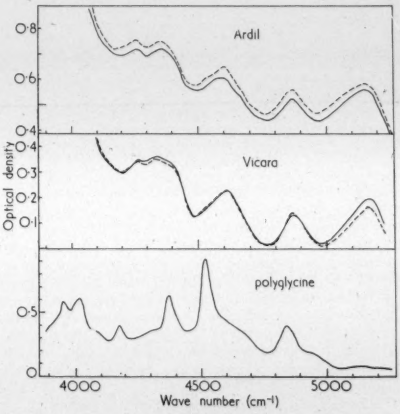


Fig. 2. Spectra of polyglycine (unoriented) and of commercial regenerated protein fibres

Full and broken lines as in Fig. 1.

prepared by casting a film from trifluoracetic acid. It is probable that it contains a proportion of small molecules. The band at 4520 cm^{-1} shows the β-configuration is present and is unusually sharp. The shoulder extending to 4700 cm^{-1} extends too far for it to be the α-component of this band and is probably a consequence of the small peptides. Possibly the shoulder on the NH combination band has the same origin.

By analogy with polythene two bands were expected arising from the CH_2 group but there is apparently only one band at 4375 cm^{-1}. The symmetrical CH_2 stretching mode is considerably weaker than the anti-symmetrical mode so that

the second band may be the very weak one at 4300 cm^{-1}. This would be consistent with the fundamental CH_2 stretching and deformation frequencies.

Wool and regenerated protein fibres.

The main difference between silks and the proteins present in wool and regenerated protein fibres is in the much larger variety of side-chains attached to the polypeptide chain in the latter in appreciable quantities. These may form cross-links of various types and favour the formation of secondary folds. It is therefore interesting to note that while the absorption bands of their polypeptide chains resemble silk and synthetic polypeptides there is an almost complete absence of dichroism (Figs. 2 and 3). The band at 4600 cm^{-1} shows that in all cases the α-configuration is predominant.

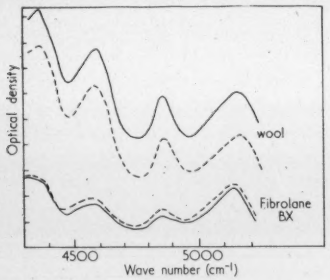


Fig. 3. Spectra of wool and of regenerated commercial casein fibre

Full and broken lines as in Fig. 1.

There is, however, in many cases an asymmetry caused by a slight shoulder towards 4520 cm^{-1}. This is usually more noticeable in the parallel position of the polarizer and indicates a certain amount of slightly oriented β-material. In the case of Fibrolane (a casein protein[21]) the orientation of the β-material is just sufficient to give detectable parallel dichroism in the NH combination band. Vicara (made from zein) appears to have rather less β-material than the other regenerated protein fibres and its spectrum most closely resembles that of wool.

The high water-uptake of these fibres is shown by the strong band at 5150 cm^{-1} but direct comparison is not possible since the measurements were not carried out under conditions of controlled humidity. Below 4400 cm^{-1} these fibres all have ill-defined bands without any distinctive features which is a consequence of the wide variety of side chains present.

Silk.

Some years ago Bath and Ellis[22] examined the spectrum of silk fibres in the overtone region, using polarized radiation. They reported marked dichroism, and refer to a perpendicular band at 1·93 μ (5180 cm^{-1}) which they ascribed to the second overtone of the C=O stretching mode. This is in the region of the band now ascribed to water, which, however, is not dichroic. We have examined silk gut which was dried at 100° C for 24 h and found that the 5150 cm^{-1} disappeared completely, with no appearance of any band which might have been overlaid by the water band. Now Bath and Ellis dried their specimens in an oven for several days at 110–115° C, and it therefore seems likely that what they observed was a band due to some product of oxidation or decomposition of silk (which is not a very heat-stable material). The thicker silk gut which we used would have very much less surface exposed to the air than the cocoon fibres used by Bath and Ellis. Their observations on the first overtone of the NH stretching mode are interesting, for this band was found to have three components showing perpendicular dichroism, which is in agreement with silk fibroin having an extended configuration. The first overtone of the NH stretching mode lies outside the region which we have examined.

The examination of silks provides a good example of the value of infra-red observations in the overtone region. Although silk gut can be sectioned to allow observations in the region of fundamental absorption bands (Ambrose and Elliott[12]), it cannot be assumed *a priori* that the structure of silk gut and cocoon silk (from the same species of moth) are identical, though the X-ray diffraction patterns suggest this.

Fig. 4 gives the spectra of several cocoon silks. Cocoon silk of Bombyx mori gives a spectrum which is indistinguishable from that of commercial silk gut. The spectra of all

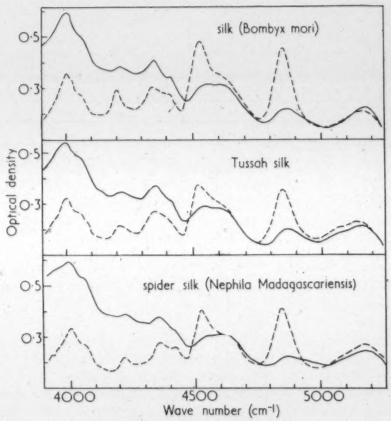


Fig. 4. Spectra of silk fibres

Full and broken lines as in Fig. 1.

three silks are notably similar to that of β-poly-*L*-alanine. In all the silks, the non-dichroic band at approximately 4600 cm^{-1} is quite prominent, and shows, in our opinion, that an appreciable fraction of the polypeptide chains in silk are in a folded configuration. There are, however, no reflexions in the X-ray diagram of the silks hitherto examined which resemble those of α-poly-*L*-alanine, and it would appear that the folded chains in silk are too amorphous to give a recognizable reflexion. Water-soluble silk made from solution in aqueous lithium bromide has a band at 4600 cm^{-1} (Toms and Elliott[23]) and gives an almost completely amorphous X-ray diagram (Ambrose, Bamford, Elliott and Hanby[24]).

The apparent high dichroism at the low-frequency end of the spectrum is not real. Because silk fibres are very thin, it is necessary to use a large number in order to obtain sufficiently strong absorption. Reflexion losses are therefore considerable, and these change with the direction of the electric vector, giving rise to what is known as "form dichroism." The effect can be reduced at any chosen wavelength by matching the refractive index of the immersion liquid suitably, but in general the effect will appear at other wavelengths, on account of dispersion. In judging dichroism, therefore, the height of a band above the local background is the best measure of intensity.

CONCLUSION

While it is clear from the above account that there are limitations to the information which can be obtained from the examination of infra-red spectra in the overtone region, it should be apparent that the method is capable of supplementing other kinds of investigation. In protein and polypeptide fibres, especially, information on chain configuration can be obtained which has hitherto not been given by other methods. There is, moreover, good reason for thinking that with the accumulation of information concerning the origin of the absorption bands more will be deduced from overtone spectra.

The small amount of extended polypeptide in the regenerated proteins seems to be well established and (though this is speculative) may mean that they are still essentially corpuscular in character. Another striking result is the fairly constant proportion of folded configuration present in the different silks examined, and in steam-stretched poly-*L*-alanine.

ACKNOWLEDGEMENTS

We wish to express our thanks to Mr. P. Robertson for the care with which he has obtained many of the spectra here shown. We are indebted to Dr. Milton Harris for specimens of Vicara. The Madagascar spider silk was obtained through the kindness of Mr. C. A. S. Grose of Courtaulds, Ltd.

REFERENCES

(1) TETLOW, K. S., MCAUSLAN, J., BRINLEY, K. J., and PRICE, W. C. *J. Sci. Instrum.*, **28**, p. 161 (1951).

(2) CZERNY, H., and TURNER, A. F. *Z. Phys.*, **61**, p. 792 (1930).

(3) ELLIOTT, A., and AMBROSE, E. J. *Nature* [*London*], **179**, p. 641 (1947).

(4) ELLIOTT, A., AMBROSE, E. J., and TEMPLE, R. B. *J. Opt. Soc. Amer.*, **38**, p. 212 (1948).

(5) ELLIOTT, A., and AMBROSE, E. J. *J. Sci. Instrum.*, **24**, p. 324 (1947).

(6) ELLIOTT, A., AMBROSE, E. J., and TEMPLE, R. B. *J. Chem. Phys.*, **16**, p. 877 (1948).

(7) JONES, E. R. S. *J. Sci. Instrum.*, **30**, p. 132 (1953).

(8) ABBOTT, N. B., and AMBROSE, E. J. *Proc. Roy. Soc. A*, **219**, p. 17 (1953).

(9) ASTBURY, W. T., and STREET, A. *Phil. Trans A*, **230**, p. 75 (1931).

(10) BAMFORD, C. H., HANBY, W. E., and HAPPEY, F. *Proc. Roy. Soc. A*, **205**, p. 30 (1951).

(11) AMBROSE, E. J., and ELLIOTT, A. *Proc. Roy. Soc. A*, **205**, p. 47 (1951).

(12) AMBROSE, E. J., and ELLIOTT, A. *Proc. Roy. Soc. A*, **206**, p. 206 (1950).

(13) GLATT, L., and ELLIS, J. W. *J. Chem. Phys.*, **16**, p. 551 (1948).

(14) PRICE, W. C., and FRASER, R. B. D. *Nature* [*London*], **170**, p. 490 (1952).

(15) ELLIOTT, A. *Textile Research J.*, **22**, p. 783 (1952).

(16) AMBROSE, E. J., and ELLIOTT, A. *Proc. Roy. Soc. A*, **208**, p. 75 (1951).

(17) ELLIOTT, A., and AMBROSE, E. J. *Disc. Faraday Soc.*, No. 9, p. 246 (1950).

(18) BAMFORD, C. H., BROWN, L., ELLIOTT, A., HANBY, W. E., and TROTTER, I. F. *Nature* [*London*], **173**, p. 27 (1954).

(19) NIELSEN, A. H., and BARKER, E. F. *Phys. Rev.*, **46**, p. 970 (1934).

(20) GLATT, L., and ELLIS, J. W. *J. Chem. Phys.*, **15**, p. 884 (1947).

(21) WORMELL, R. L. *J. Textile Inst.*, **44**, p. 258 (1953).

(22) BATH, J. D., and ELLIS, J. W. *J. Phys. Chem.*, **45**, p. 204 (1941).

(23) TOMS, B. A., and ELLIOTT, A. *Nature* [*London*], **169**, p. 877 (1952).

(24) AMBROSE, E. J., BAMFORD, C. H., ELLIOTT, A., and HANBY, W. E. *Nature* [*London*], **167**, p. 264 (1951).

Further tests on the stability of analytical weights in chemical laboratories

By P. H. Bigg, B.Sc., and F. H. Burch, F.B.H.I., National Physical Laboratory, Teddington, Middlesex

[*Paper first received* 11 *May, and in final form* 14 *June*, 1954]

The stability, in the corrosive atmospheres of chemical laboratories, of brass weights coated electrolytically with nominal thicknesses of 13 μ and 25 μ of tin-nickel alloy (65% Sn, 35% Ni) has been compared with that of weights previously tested[1] under similar conditions. As regards appearance after exposure, the 25 μ coating was the best, and was comparable with weights of highly-polished stainless steel save in the severest conditions. All the alloy-coated weights showed about the same stability of mass as good quality commercially produced weights of austenitic stainless steel (25% Cr, 20% Ni) and rhodium-plated brass weights. Tin-nickel alloy plating does not possess any undesirable magnetic properties. In highly corrosive atmospheres, non-magnetic nickel-chromium (80% Ni, 20% Cr) is a little less resistant than the materials already mentioned. Of the small sheet metal weights of various materials tested, those of austenitic stainless steel were definitely the best. They were followed, in order of merit, by zirconium, tantalum and aluminium, and finally, titanium.

In some experiments reported in 1950[1] analytical weights of various modern materials were subjected to accelerated corrosion and stability tests in representative chemical laboratories. It was found that as regards stability of mass there was little to choose between weights of austenitic stainless steel (25% Cr, 20% Ni), non-magnetic nickel-chromium (80% Ni, 20% Cr) and chromium-, rhodium- and platinum-plated brass weights having a nominal thickness of plating of about 15 μ (0·0006 in.) or more, and that some weights of specially highly-polished 25/20 stainless steel were slightly superior.

More recently, details of an attractive new plating material, namely, tin-nickel alloy (65% Sn, 35% Ni), have been published by the Tin Research Institute.[2,3] Experiments carried out by the Institute indicated that this plating resists tarnishing in severely polluted atmospheres, being unaffected by sulphur dioxide or hydrogen sulphide, and differs from nickel-plating in not being subject to "fogging" under ordinary atmospheric conditions. It is resistant to corrosion at ordinary temperatures by alkalis, neutral solutions, and nitric acid, and by other acids at pH values greater than 1·2. Deposition is easy to control, and occurs fairly uniformly irrespective of the shape of the plated article. The deposit is hard, and a light buffing gives it a very pleasant lustrous appearance.

On account of these outstanding qualities, tests on tin-nickel plated brass weights were put in hand, broadly on the same lines as those previously reported. Three variants were tested, namely, brass coated with a nominal thickness of (*a*) 13 μ of tin-nickel alloy; (*b*) 25 μ of tin-nickel alloy; (*c*) 13 μ of tin-nickel alloy deposited on 13 μ of copper.

According to the Tin Research Institute, a good quality coating of 13 μ (0·0005 in.) of tin-nickel alloy on brass is enough to keep porosity within tolerable limits for normal service, and a coating 25 μ (0·001 in.) thick should be capable of good service under severe conditions. For this reason the variants (*a*) and (*b*) were selected for trial in the atmosphere of chemical laboratories. Variant (*c*) was introduced as it was thought that the copper might help to cover pin-holes, and it was of interest to compare the performance of variant (*b*) with that of variant (*c*). Weights of 25/20 austenitic stainless steel and of 80/20 nickel-chromium were included to serve as a basis of reference and comparison with the previous results. Some good quality rhodium-plated weights were also included, as representing plated weights which are normally available; they were taken from a firm's ordinary production. The tests were extended to include sheet metal weights of 25/20 austenitic stainless steel, aluminium, tantalum, titanium and zirconium. The stainless steel weights were included mainly to provide a basis of comparison with the various materials used in the earlier tests. It was considered unnecessary to supplement them by similar nickel-chromium weights. Aluminium is used in this country for very small denominations only, but larger aluminium weights are used abroad. Tantalum has also been used abroad. It seemed desirable to include titanium and zirconium, two new metals of low density. The use of the latter for fractional weights has been recommended by Thornton.[4] All the samples used were those readily available, and were regarded as representative in quality of those which might be used industrially.

Details of all the weights exposed are given in Table 1. Except for the stainless steel cylinders, which were the three used in the previous experiments[1] and since repolished, six representatives of each type of weight were provided. The nominal thicknesses of plating quoted in Table 1 for the tin-nickel alloy were a little less than those calculated from the average changes in mass due to plating at the Tin Research Institute, but direct measurements under high magnification of representative polished sections embedded in bakelite indicated that the actual average thicknesses were rather less than nominal. In each specimen measured the thickness of plating was greater on the flank than on the top of the weight, and was slightly greater still at the curved edge. The thicknesses given for the rhodium and silver coatings were the average values deduced from the increases in weight due to the plating, according to the manufacturer's data. The measured thickness both on the flank and at the edge of a single sample was found to agree closely with the deduced thickness.

In view of the effect of the use of corrodible material in the adjustment of screw knob weights [see ref. (1), Table 5], care was taken to ensure that only resistant material was used in the present instance.

The weights were divided equally between three laboratories so that there were two of each type (but only one

Table 1. *Details of weights exposed, and average changes found*

100 g weights	Form of weights	Stability in preliminary period (January/March) under good conditions (unit 0·01 mg)			Mean changes* (unit 0·01 mg) during exposure in chemical laboratories				
		Mean change excluding sign	Mean change including sign	Greatest change	First month	Second month	Third month	First 2 months	All 3 months
Tin-nickel alloy plated weights	Screw knob, tall form								
13 μ tin-nickel alloy on brass		0·9	−0·5	−2	6	9	10	17	28
25 μ tin-nickel alloy on brass		1·7	−0·2	−3	4	8	10	14	24
13 μ tin-nickel alloy on 13 μ copper on brass		1·1	+1·0	+3	5	11	9	16	22
Rhodium plated weights from normal production 0·8 μ rhodium on 25 μ silver on brass	Screw knob, tall form	1·6	+0·8	−2	9	9	6	17	22
Stainless steel weights, 25% Cr/20% Ni, from normal production	Screw knob, tall form	1·5	−0·7	−4	6	13	4	17	20
Stainless steel cylinders, 25% Cr/20% Ni, produced at N.P.L., with specially high polish	Plain cylinder, approx. 25 mm × 25 mm	—	—	—	10	6	2	8	12
Non-magnetic nickel-chromium weights, 80% Ni/20% Cr from normal production	One piece knob weights, tall form	1·1	0·0	−2	25	14	10	38	46

Sheet metal weights	Details of weights				Mean changes* (unit 0·001 mg)				
	Value (g)	Area of upper surface (cm²)	Density (g/ml)		First month	Second month	Third month	First 2 months	All 3 months
Aluminium	0·1	3·61	2·7	*A*†	2·8	6·1	5·3	8·9	13·3
				B‡	1·0	2·3	1·9	3·3	4·7
Titanium	0·5	3·24	4·5	*A*	5·9	12·4	3·7	18·6	18·8
				B	1·3	2·7	0·8	4·1	4·2
Zirconium	0·5	2·47	6·5	*A*	2·4	3·4	4·7	5·1	10·4
				B	0·4	0·5	0·7	0·8	1·6
Stainless steel, 25% Cr/20% Ni	1·0	2·59	7·8	*A*	1·5	0·8	1·9	2·3	4·3
				B	0·2	0·1	0·2	0·3	0·5
Tantalum	0·5	3·24	16·6	*A*	1·9	6·8	7·1	8·6	15·5
				B	0·1	0·4	0·4	0·5	0·9

* Based on adjusted values (see under *Changes of mass*).
† *A* = mean change per 1 cm² of upper surface.
‡ *B* = *A*/density.

stainless steel cylinder) in each laboratory. They were placed symmetrically on acid-free tissue paper in the specially designed ventilated anodized-aluminium containers previously described,[1] and suspended slightly above head level.

The values of the 100 g weights were determined on several occasions between their receipt and first exposure. They all proved satisfactorily stable in relation to the accuracy of an individual weighing, ±0·02 mg (see Table 1). The values of the fractional weights were determined with an accuracy generally within ±0·003 mg. All the weights were re-weighed at the N.P.L. after an exposure of one month, weighed again after a second month's exposure, and re-weighed after a third. During transportation before and after weighing the 100 g weights were held in specially designed boxes by gentle pressure on the tops of the knobs so as to minimize interference with the condition of the exposed surfaces. The fractional weights did not need to be secured during transit. In accordance with normal procedure all the weights were lightly dusted with a camel-hair brush immediately before weighing, and were subjected to no other conditioning.

OBSERVED RESULTS

Appearance. In all three laboratories the appearance of the weights plated with 13 μ of tin-nickel alloy was not so good as that of those plated with 25 μ tin-nickel alloy, or 13 μ copper plus 13 μ tin-nickel alloy. The weights plated with 25 μ tin-nickel alloy were better than those with 13 μ copper plus 13 μ tin-nickel alloy in the severest conditions only, and except in these conditions they were comparable

with the highly-polished stainless steel cylindrical weights. The latter were little affected, even in the most severe conditions. A light film remained on all the weights after brushing. On the tin-nickel plated weights it was no worse than on any of the others, and on removal by light wiping with a soft wash-leather after the completion of the experiments a lustrous surface was revealed. The rhodium-plated weights early developed white misty patches in the most severe conditions, and similar patches developed later in the less severe conditions, with "hairline" corrosion over the exposed surfaces of all the six weights. The nickel-chromium and stainless steel weights from ordinary production were little affected in the least severe conditions, but in the other laboratories, especially in those with the most severe conditions, they showed extensive spotting and some discoloration.

Very little change could be discerned in the appearance of any of the sheet metal weights apart from those of zirconium and aluminium which, in the laboratories with the more severe conditions, became slightly dull.

Changes of mass. Table 1 gives the mean changes of mass for each type of weight. Apart from the stainless steel cylinders, of which there was only one in each of the three laboratories, there were six pairs of 100 g weights and five pairs of sheet metal weights in each laboratory. Hence for the five exposure periods indicated in Table 1 there were ninety pairs of observed changes for the 100 g weights and seventy-five pairs for the sheet metal weights. In regard to the 100 g weights, after allowing twice 0·02 mg on account of possible weighing errors the individual values of a pair agreed within 10% except in nine pairs. In only one instance (nickel-chromium) did the difference exceed 40% of the larger change. The corresponding figures for the sheet metal weights were thirteen of the seventy-five values, twice 0·003 mg having been allowed for possible weighing errors, and in only two instances (both tantalum) did the difference exceed 30% of the larger change. As in the previous work, in order to avoid giving undue weight in the final mean values to the results from the more severely corrosive laboratories, the observed changes in individual weights in a given laboratory were adjusted on the basis of their sum, and the corresponding sums in the other laboratories, to bring the total change in each laboratory to the same level. These adjustments naturally destroyed the simple numerical relationships between the changes for different exposure periods and sums of periods, such as would otherwise have appeared in Table 1. The relative severity of conditions in the three laboratories, in terms of the average changes of mass, is indicated in Table 2.

Table 2. *Relative severity of conditions in the three laboratories*

	Mean change (mg) during					
	first exposure		*second exposure*		*third exposure*	
Laboratory	*100 g weights*	*sheet metal weights*	*100 g weights*	*sheet metal weights*	*100 g weights*	*sheet metal weights*
A	0·01	0·001	0·02	0·003	0·02	0·003
D	0·02	0·003	0·01	0·005	0·04	0·007
G	0·09	0·005	0·09	0·011	0·03	0·005

The laboratories *A* and *D* were the laboratories designated *A* and *D* in the previous paper.[1] A continuous record of the relative humidity in each laboratory revealed that the level, 70%, above which the tendency towards corrosion markedly increases[5] was exceeded on a number of occasions. The humidity in the laboratory *G* was by far the highest as regards both the frequency and the total duration of the high values.

As regards change of mass there was little to choose between the three variants of tin-nickel alloy plating included in the test. They were all as good as the stainless steel and rhodium-plated weights from normal production which were exposed with them, and at first they even tended to show slight superiority. However, before drawing general conclusions in comparison with other varieties of weights the results of previous experiments should be recalled.

Ordinarily-finished screw knob weights of stainless steel and of 80/20 nickel-chromium, and the platinum-, rhodium- and chromium-plated weights having a total thickness of about 15 μ or more of plating, were then regarded as constituting a single good category with but little to choose between the individual kinds. In this group the 80/20 nickel-chromium weights were not among the most stable. A few highly-polished cylinders of 25/20 stainless steel, exposed during one period only, changed the least of all. These earlier results are borne out in the present experiments. In the very corrosive atmospheres encountered the highly-polished stainless steel cylinders were again the most stable on the whole, though not initially, and the (ordinary) nickel-chromium weights were not so good as the stainless steel weights from normal production, or the rhodium-plated weights. The stability of tin-nickel alloy coatings of the quality provided therefore appears to be at least as high as that of most of the materials previously studied. It must not be overlooked, of course, that weights plated with this material are not likely to be specially immune from the very occasional (and unexplained) instability noted at the N.P.L. in regard to other plated weights.[1]

Through the courtesy of Baird and Tatlock (London) Ltd. it is possible to compare the rates of change of some tin-nickel alloy coated weights and an 80/20 nickel-chromium weight exposed by them over a period of six months, with those obtained in the three laboratories from which the results have been reported above. To relate them to the total changes found by the N.P.L. over the full three months exposure, their values have been halved.

Table 3. *Comparison of results from N.P.L. and Baird and Tatlock (London) Ltd.*

(unit 1 mg)

	Lab. A	*Lab. D*	*Lab. G*	*B.T.L.*
13 μ Sn–Ni	0·12	0·22	0·34	0·0
25 μ Sn–Ni	0·13	0·12	0·36	$0·1_5$
13 μ Sn–Ni on 13 μ Cu	0·11	0·12	0·47	$0·1_5$
80/20 Ni Cr	0·15	0·30	1·03	0·1

The results are given in Table 3, which shows that the rate of change in the laboratory of Baird and Tatlock (London) Ltd. was about the same as in the two laboratories with less severe conditions which housed the N.P.L. weights, but suggests also that nickel-chromium is not appreciably inferior except under severely corrosive conditions.

Turning now to the sheet metal weights, for each pair of samples and each exposure period two mean values are given in Table 1. Those designated *A*, namely, the change in mass (μg) per unit area of upper surface, indicate the relative efficiency of the materials when used for all denominations but the lowest, because for practical reasons most sheet metal weights of a common denomination have about the same

area, density differences between materials being taken up on thickness. On the basis of these results the 25/20 stainless steel was quite clearly superior to all the other materials, which took, broadly, the following order of merit: zirconium, aluminium and tantalum, titanium. The values designated *B* were obtained by dividing the *A* values by the density of the material, and are mainly of interest in connexion with the lower denominations of weights, i.e. those in which it is undesirable to reduce the thickness sufficiently to arrive at a convenient surface area. The 25/20 stainless steel was the best on this basis also, followed, in order of decreasing merit, by tantalum, then zirconium, with aluminium and titanium roughly equal in the lowest place. The high density of tantalum makes it inappropriate for the smallest denominations, and zirconium has insufficient advantage in density to offset its inferiority in resistance. Titanium is 70% more dense than aluminium, to which it did not prove superior in other respects. Therefore, of the materials tested in sheet-metal form, stainless steel emerges definitely the best, while aluminium has not been displaced as the most suitable material for the smallest fractions. It should be noted, however, that it is now normal practice to supply fractions down to and including 0·01 g, the lowest denomination recommended by the N.P.L. for fractional weights, in materials of density about 8 g/ml. 10 mg weights of this density are of reasonable size, and therefore there would seem to be no point, in ordinary circumstances, in using aluminium. Fractional weights and riders formed from wire generally have smaller surface areas than sheet metal weights of the same denominations; therefore relatively smaller changes of mass may be expected to occur in wire weights than in those of sheet metal.

Table 4 brings together results derived from the 1953 and 1950 tests, and also the changes shown by nominally the same material in different forms and surface conditions. The values are related to unit area of exposed surface (i.e. excluding supporting surface) for comparability. The rates of change in 1953 and 1950 are roughly of the same magnitude, and are in broad agreement in regard to the relative resistance associated with the different types of finish. It is also of interest to note that for the stainless steel the change experienced by the sheet metal weights is closer to that for the highly-polished cylinders than for the weights from ordinary production. This might have been due to the surface condition (the sheet metal weights were made from rolled strip). However, the sheet metal weights rested on the floor of the container, and it was noticeable on the 100 g weights that attack was less severe there than a little higher up. Perhaps the conditions there were more favourable. In Table 4 the superiority of 25/20 stainless steel over 80/20 nickel-chromium in highly corrosive atmospheres is again apparent.

Table 4. *Comparison of rates of increase of mass expressed in μg per cm² per month*

		25/20 stainless steel	80/20 nickel-chromium
1953	Screw knob weights from normal production	2·7	–
	Cylinders, highly polished	1·6	6
	Sheet metal weights	1·5	–
1950	Screw knob weights from normal production	–	5
	Cylinders, turned	3	–
	Cylinders, highly polished	1	4

MAGNETIC PROPERTIES OF THE TIN-NICKEL ALLOY PLATING

In order to check that the tin-nickel alloy does not possess magnetic properties which would render it unsuitable for good quality weights, two 100 g blanks of electrolytic copper plated respectively with 13 μ and 25 μ of tin-nickel alloy, were magnetized in a field of approximately 2000 oersteds. In neither instance was the apparent permeability greater than 1·0005, or the residual magnetic moment per unit volume greater than 0·002 c.g.s. units. These values are very satisfactorily low (see Gould[6]), and indicate that magnetically, tin-nickel alloy is completely satisfactory for plating weights.

CONCLUSIONS

(1) Brass weights coated with tin-nickel alloy plating nominally 13 μ and 25 μ in thickness, and with 13 μ tin-nickel alloy on 13 μ of copper, have proved about equally resistant to change of mass in the corrosive atmospheres of three chemical laboratories, and comparable with austenitic stainless steel (25% Cr, 20% Ni), and with platinum-, rhodium- or chromium-plated brass weights having the same minimum overall thickness of plating. Judged by appearance, the weights coated with a total of 25 μ of deposit (either tin-nickel alone, or 13 μ of tin-nickel on 13 μ of copper) showed greater resistance than those with 13 μ of tin-nickel deposited direct on brass, and in the most severe conditions the coating consisting solely of 25 μ of tin-nickel alloy proved the best. Unlike nickel, the alloy does not appear to be liable to "fogging".

(2) Tin-nickel alloy plating does not possess any undesirable magnetic properties.

(3) Tin-nickel alloy plating has a very pleasant appearance, and according to published accounts the plating process is easily controlled. Therefore it seems an attractive material for coating weights.

(4) In highly corrosive atmospheres, weights of non-magnetic nickel-chromium (80% Ni, 20% Cr) are a little less resistant than those mentioned in conclusion (1). This does not detract from their merit in ordinary circumstances.

(5) In comparisons of sheet metal fractional weights of austenitic stainless steel (25% Cr, 20% Ni), aluminium, titanium, zirconium and tantalum, the stainless steel proved to be the best.

Note added in proof. The above paper was primarily concerned with the resistance to corrosion of weights nominally plated to a given thickness, and the details of the actual plating thickness were regarded as of minor interest. Baird and Tatlock (London) Ltd. have more recently directed detailed attention to the variation of thickness of plating on a specimen of the tin-nickel weights having a nominal thickness of 25 μ. They have confirmed the general conclusions given above and have found that on the top of the weight the actual thickness diminished from about 23 μ at the edge to about 7 μ near the centre. While serving to indicate further the effectiveness of tin-nickel alloy as a protective coating on weights, these measurements also demonstrate the need for specifying a high average plating thickness.

ACKNOWLEDGEMENTS

The work described has been carried out as part of the research programme of the N.P.L., and this paper is published by permission of the Director of the Laboratory. The N.P.L. is grateful to the Research and Development Division of Baird and Tatlock (London) Ltd. for drawing its attention to the development of tin-nickel alloy plating, and for

furnishing later the results of tests made in their laboratories. Thanks are also due to Baird and Tatlock (London) Ltd., and to Stanton Instruments Ltd., for supplying many of the weights used. The co-operation of the directors and staffs of the laboratories of the Distillers Co. Ltd., and the Fuel Research Station, where two of the sets of weights were housed, is appreciated.

The authors desire to acknowledge the assistance of a number of their colleagues, and the help received from Miss D. Oxley who was responsible for the weighings and made most of the computations, and Miss J. M. Holloway and Miss H. M. Richardson who carried out the weighings.

REFERENCES

(1) Bigg, P. H., and Burch, F. H. *Brit. J. Appl. Phys.*, **2**, p. 126 (1951).

(2) Parkinson, N. *J. Electrodep. Tech. Soc.*, **27**, p. 129 (1951).

(3) Britton, S. C., and Angles, R. M. *J. Electrodep. Tech. Soc.*, **27**, p. 293 (1951).

(4) Thornton, W. M. *J. Franklin Inst.*, **250**, p. 39 (1950).

(5) Vernon, W. H. J. *Trans Faraday Soc.*, **23**, p. 159 (1927); **31**, p. 1678 (1935).

(6) Gould, F. A. *J. Sci. Instrum.*, **23**, p. 124 (1946).

The use of semiconductors in thermoelectric refrigeration

By H. J. Goldsmid, B.Sc., and R. W. Douglas, B.Sc., F.S.G.T., F.Inst.P., Research Laboratories, The General Electric Co. Ltd., Wembley, Middlesex

[*Paper received* 6 *July*, 1954]

In the past the possibility of thermoelectric refrigeration has been considered, but all attempts to produce a practical refrigerator have failed owing to lack of suitable thermocouple materials. In this paper it is proposed that semiconductors should be used and the factors governing their selection are discussed. It is concluded that the semiconductors should be chosen with high mean atomic weights and that they should be prepared with thermoelectric powers lying between 200 and 300 μV. °C^{-1}. Preliminary experiments have led to the production of a thermocouple consisting of bismuth telluride, Bi_2Te_3, and bismuth, capable of maintaining 26° C of cooling.

In 1834 Peltier discovered that there was a heating or cooling effect, quite apart from the ordinary resistance heating, whenever an electric current flowed through a junction between two different metals. Several attempts have been made to devise a thermoelectric refrigerator using the Peltier cooling effect. All these attempts have been unsuccessful because no materials were available with the requisite properties for use as thermocouple elements.

The difficulties arise from the fact that the Peltier cooling is generally much less than the Joule heating. If the electrical resistance of the thermocouples is decreased by constructing them of elements of short length and large cross-section area, it is found that the heat conducted from the hot junctions to the cold junctions becomes excessive. The problem is to find a material with a high thermoelectric power and a high ratio of electrical conductivity to thermal conductivity.

The ratio of electrical to thermal conductivity is approximately constant for all *metals* at any given temperature. The best metals, from the point of view of thermoelectric refrigeration, are therefore those with the highest thermoelectric powers. These are antimony and bismuth and their alloys. Even with such materials as these, however, the thermoelectric powers are not sufficient. The largest temperature difference that has been reported is only 10° C, which was obtained from a junction between two bismuth alloys.[1] It thus seems that the use of metallic thermocouples holds little promise.

In the past few years great interest has been centred on the properties of semiconductors. Although, in general, the ratio of electrical to thermal conductivity in these materials is much less than it would be in a metal, the thermoelectric power may be much greater, and it is therefore worthwhile considering the use of semiconductors in thermoelectric refrigerators.

THEORY OF THERMOELECTRIC REFRIGERATION

The theory of thermoelectric refrigeration has been given by Altenkirch[2]; his results are summarized in this section. Altenkirch's model consisted of a single thermocouple of two elements with a non-resistive junction between them. The other ends of the elements were assumed to be connected to a battery and maintained at a set temperature. He assumed elements of uniform cross-section, but Gehlhoff and others[3] have shown that no advantage is to be gained from the use of elements of other forms. The latter authors also pointed out the great importance of good electrical contacts at the junctions.

The two quantities of interest are the maximum temperature difference that can possibly be reached using a given thermocouple and the coefficient of performance for any lesser temperature difference. The coefficient of performance is defined as

$$\phi = \text{refrigerating effect/work expended}$$

It should be noted that ϕ can be greater than unity. In fact, for an ideal thermodynamic machine,

$$\phi = T_2/(T_1 - T_2)$$

where T_1 = temperature of sink (°K).
T_2 = temperature of source (°K).

The properties of importance for the thermocouple elements are

η = thermoelectric power (V. °C^{-1}).
λ = thermal conductivity (W. cm^{-1}. °C^{-1});
σ = electrical conductivity (Ω^{-1}. cm^{-1});
l = length (cm);
S = cross-section area (cm^2).

The subscripts a and b will be used to denote the two materials. A factor θ will be defined as

$$\theta = \frac{|\eta_a - \eta_b|}{\sqrt{(\lambda_a/\sigma_a)} + \sqrt{(\lambda_b/\sigma_b)}} \frac{(T_1 + T_2)^{\frac{1}{2}}}{2}$$

It can be shown that, for the coefficient of performance ϕ, to be a maximum, it is necessary that the dimensions of the thermocouple elements obey the relation

$$\frac{l_a S_b}{l_b S_a} = \left(\frac{\lambda_a \sigma_a}{\lambda_b \sigma_b}\right)^{\frac{1}{2}} \tag{1}$$

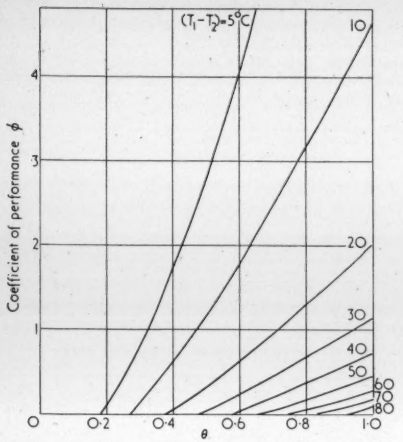


Fig. 1. Variation of coefficient of performance (ϕ) with figure of merit (θ) for given temperature differences ($T_1 - T_2$). Hot junction assumed to be at 300° K

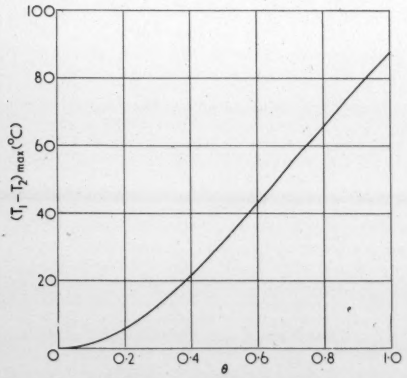


Fig. 2. Variation of maximum temperature difference $(T_1 - T_2)_{max}$ with figure of merit (θ). Hot junction assumed to be at 300° K

This being so, the coefficient of performance for a given temperature difference, when the optimum voltage is applied, is

$$\phi = \frac{T_2\sqrt{(1+\theta^2)} - T_1}{(T_1 - T_2)[1 + \sqrt{(1+\theta^2)}]} \tag{2}$$

Fig. 1 shows the variation of the coefficient of performance with θ and the temperature difference. It is clear that θ represents a figure of merit for the particular thermocouple materials, since ϕ increases as θ increases for a given temperature difference. The maximum temperature difference is reached when $\phi = 0$, and then

$$\left(\frac{T_1}{T_2}\right)_{max} = \sqrt{(1+\theta^2)} \tag{3}$$

Fig. 2 shows the way in which the maximum temperature difference increases with θ.

The reason for the lack of success in the past has been the lowness of the values of the figure of merit, θ, associated with all known couples.

THERMOELECTRIC PROPERTIES OF SEMICONDUCTORS

We shall now extend this theory by discussing the factors which govern the value of θ for semiconductor junctions. Both the thermoelectric power and the electrical conductivity of a semiconductor depend on the concentration of current carriers.[4] However, the thermal conductivity is determined not only by an electronic component but also by a lattice component. It may be almost independent of the carrier concentration when the major contribution to the conduction of heat comes from the elastic vibrations of the lattice as is the case with germanium.

The sign of the thermoelectric power in a semiconductor is the same as the sign of the effective charge on the majority carriers. Thus for the thermoelectric power of a thermocouple to be large it should be composed of an n-type and a p-type semiconductor. In the n-type semiconductor the conductivity is given by

$$\sigma_n = ne\mu_n \tag{4a}$$

where n = concentration of free electrons (cm^{-3});
e = electronic charge (coulomb);
μ_n = mobility of electrons ($cm^2 . V^{-1} . s^{-1}$);

similarly in the p-type semiconductor

$$\sigma_p = pe\mu_p \tag{4b}$$

where p = concentration of positive holes (cm^{-3});
μ_p = mobility of holes ($cm^{-2} . V^{-1} . s^{-1}$).

A preliminary calculation shows that at the optimum positions for the Fermi level there is a partial degree of degeneracy in the semiconductors. The expressions for the carrier concentrations are therefore

$$\left.\begin{aligned} n &= 2\left(1{\cdot}6\times10^{-12}\times\frac{2\pi kT}{h^2}\right)^{\frac{3}{2}} m_n^{*\frac{3}{2}}\left[\frac{1}{0{\cdot}25 + \exp(\epsilon_n/kT)}\right] \\ p &= 2\left(1{\cdot}6\times10^{-12}\times\frac{2\pi kT}{h^2}\right)^{\frac{3}{2}} m_p^{*\frac{3}{2}}\left[\frac{1}{0{\cdot}25 + \exp(\epsilon_p/kT)}\right] \end{aligned}\right\} \tag{5}$$

using the approximation to the Fermi function suggested by Ehrenberg,[5] where

k = Boltzmann's constant (electron volt. °C^{-1});
h = Planck's constant (electron volt. s);

m_n^* = effective mass of the free-electrons in the n-type semiconductor (g);

m_p^* = effective mass of the positive holes in the p-type semiconductor (g);

ϵ_n = difference of energy between the bottom of the conduction band and the Fermi level in the n-type semiconductor (electron volt);

ϵ_p = difference of energy between the Fermi level and the top of the valence band in the p-type semiconductor (electron volt).

Further, the thermoelectric powers are[6]

$$\left.\begin{aligned}\eta_n &= -\left(I_n + \frac{\epsilon_n}{T}\right)\\ \eta_p &= \left(I_p + \frac{\epsilon_p}{T}\right)\end{aligned}\right\} \quad (6)$$

where

$$I_n = \frac{\int_0^\infty \tau_n E^{\frac{1}{2}} \exp\left(-\frac{E}{kT}\right) dE}{\int_0^\infty \tau_n E^{\frac{1}{2}} \exp\left(-\frac{E}{kT}\right) dE}$$

$$I_p = \frac{\int_0^\infty \tau_p E^{\frac{1}{2}} \exp\left(-\frac{E}{kT}\right) dE}{\int_0^\infty \tau_p E^{\frac{1}{2}} \exp\left(-\frac{E}{kT}\right) dE}$$

where τ_n and τ_p are the respective relaxation times for n-type and p-type semiconductors, and E is the energy of an electron. From equations (4), (5) and (6), substituting the suffixes n and p for a and b, the figure of merit, ϑ, for the couple is

$$\theta = \frac{\left(I_n + I_p + \frac{\epsilon_n + \epsilon_p}{T}\right)\left[2e\left(1{\cdot}6 \times 10^{-12} \times \frac{2\pi kT}{h^2}\right)^{\frac{3}{2}}\left(\frac{T_1 + T_2}{2}\right)\right]^{\frac{1}{2}}}{\left\{\frac{\lambda_n}{\mu_n m_n^{*\frac{3}{2}}}\left[0{\cdot}25 + \exp\left(\frac{\epsilon_n}{kT}\right)\right]\right\}^{\frac{1}{2}} + \left\{\frac{\lambda_p}{\mu_p m_p^{*\frac{3}{2}}}\left[0{\cdot}25 + \exp\left(\frac{\epsilon_p}{kT}\right)\right]\right\}^{\frac{1}{2}}} \quad (7)$$

θ is at a maximum with respect to ϵ_n and ϵ_p when $\partial\theta/\partial\epsilon_n = \partial\theta/\partial\epsilon_p = 0$. In the general case the calculation of this condition is complicated by the fact that I_n, μ_n and λ_n depend on ϵ_n, and I_t, μ_p and λ_p depend on ϵ_p. It will, however, be assumed here that the variations of I_n, I_p, μ_n, μ_p, λ_n and λ_p are sufficiently slow to be neglected and then the conditions for θ to be a maximum are

$$\left.\begin{aligned}\frac{\lambda_n}{\mu_n m_n^{*\frac{3}{2}}} \frac{\exp\left(\frac{2\epsilon_n}{kT}\right)}{\left[0{\cdot}25 + \exp\left(\frac{\epsilon_n}{kT}\right)\right]} &= \frac{\lambda_p}{\mu_p m_p^{*\frac{3}{2}}} \frac{\exp\left(\frac{2\epsilon_p}{kT}\right)}{\left[0{\cdot}25 + \exp\left(\frac{\epsilon_p}{kT}\right)\right]}\\ \text{and}\\ \frac{I_n + I_p}{k} + \frac{\epsilon_n + \epsilon_p}{kT} &= 4 + \frac{1}{2}\left[\exp\left(\frac{-\epsilon_n}{kT}\right) + \exp\left(\frac{-\epsilon_p}{kT}\right)\right]\end{aligned}\right\} \quad (8)$$

It is possible to simplify these expressions if it is assumed that the couple is made from n-type and p-type specimens of the same semiconducting material. Then $\lambda_n = \lambda_p = \lambda$ if the lattice component of the thermal conductivity is the dominating factor. The following approximation will be used in order to simplify equations (8):

$$\frac{\lambda_n}{\mu_n m_n^{*\frac{3}{2}}} \simeq \frac{\lambda_p}{\mu_p m_p^{*\frac{3}{2}}}$$

This implies that μ_n/μ_p should be approximately equal to $(m_p^*/m_n^*)^{\frac{3}{2}}$.

In fact $\mu_n/\mu_p = (m_p^*/m_n^*)^{\frac{3}{2}}$ for polar-type lattice scattering,[7] but for scattering from ionized impurity centres,[8] $\mu_n/\mu_p = (m_p^*/m_n^*)^{\frac{1}{2}}$ and for covalent-type lattice scattering[9] $\mu_n/\mu_p = (\epsilon_p'/\epsilon_n')^2(m_n^*/m_n^*)^{\frac{1}{2}}$, where ϵ_n' and ϵ_p' are the rates of change of band edges with lattice parameter. The approximation will therefore only hold good in general so long as m_p^* and m_n^* do not differ too considerably from each other. Then equations (8) become:

$$\left.\begin{aligned}\epsilon_n &= \epsilon_p = \epsilon\\ 2\frac{\epsilon}{kT} - \exp\left(\frac{-\epsilon}{kT}\right) &= 4 - \frac{I_n + I_p}{k}\end{aligned}\right\} \quad (9)$$

Now

$I_n + I_p = 8k$ for ionized impurity scattering ($\tau \propto E^{\frac{1}{2}}$);

$= 5k$ for polar-type lattice scattering (τ independent of E);

$= 4k$ for covalent-type lattice scattering ($\tau \propto E^{-\frac{1}{2}}$).

These values are only strictly true for non-degenerate cases, but are approximately correct for the partial degree of degeneracy considered here.

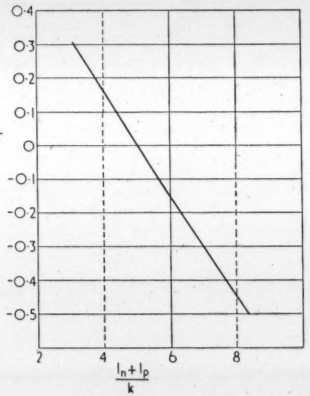


Fig. 3. Graph of $\frac{\epsilon_n}{kT} = \frac{\epsilon_p}{kT} = \frac{\epsilon}{kT}$ against $\frac{(I_n + I_p)}{k}$ for θ to be a maximum

Fig. 3 shows how ϵ/kT would vary with $(I_n + I_p)/k$ for θ to be a maximum according to the relations (9). From this diagram the optimum values of ϵ/kT are

$\epsilon/kT = -0{\cdot}45$ for $(I_n + I_p)/k = 8$

$= 0$ for $(I_n + I_p)/k = 5$

$= 0{\cdot}16$ for $(I_n + I_p)/k = 4$

from equation (6), therefore, the optimum thermoelectric power for each semiconductor is

$$\left.\begin{aligned}\eta &= k\,(4-0{\cdot}45) = 307\ \mu\text{V.}\,^{\circ}\text{C}^{-1} \text{ for ionized impurity scattering}\\ &= k\,(3) \quad\ = 260\ \mu\text{V.}\,^{\circ}\text{C}^{-1} \text{ for polar-type lattice scattering}\\ &= k\,(2\times 0{\cdot}16) = 188\ \mu\text{V.}\,^{\circ}\text{C}^{-1} \text{ for covalent-type lattice scattering}\end{aligned}\right\} \quad (10)$$

It therefore seems apparent that the materials should be prepared with thermoelectric powers lying between about 200 and 300 μV. $^{\circ}$C^{-1}. The exact value should be determined empirically in each case, in view of the approximations which have been made. In particular the assumption that the thermal conductivity of the semiconductor is almost independent of the carrier concentration may not be admissible for some of the most favourable cases where the ratio of electrical to thermal conductivity might approach the value predicted by the Wiedemann–Franz law for metals. The effect of this would be to increase the optimum value of the thermoelectric power.

A further important fact arises from equation (7), namely that the figure of merit, θ, increases with the value of $(\mu/\lambda)m^{*\frac{3}{2}}$ for each of the materials. It is therefore desirable that the semiconductors should be chosen with the highest ratio of carrier mobility to thermal conductivity, and that the effective mass of the carriers should be as great as possible. It has recently been shown[10] that the ratio of carrier mobility (limited by lattice scattering only) to thermal conductivity increases rapidly with the atomic weight of the semiconductor (or the mean atomic weight in the case of compounds). Further, the mobility, as limited by impurity scattering only, increases with the dielectric constant, and the latter quantity also rises, in general, with the atomic weight. It is thus deduced that the semiconductors for use in thermoelectric refrigerators should be chosen with the highest mean atomic weights, where this is consistent with a large effective mass of the charge carriers.

APPLICATION OF THE THEORY TO GERMANIUM

Experimental data have been published for germanium which enable the theoretical predictions to be examined quantitatively. Fig. 4 shows how the thermoelectric power varies with the resistivity according to the results of Middleton and Scanlon.[11] In the case of the *n*-type material the experimental curve has been extended using the theoretical formula for the thermoelectric power and the measured values for the resistivity and the mobility.[12]

At the optimum value of the thermoelectric power predicted from equations (10) for any semiconductor the product of thermoelectric power and the root of electrical conductivity ($\eta\sqrt{\sigma}$) should reach a maximum since, in view of the assumptions that have been made, $\lambda_n = \lambda_p$, $\sigma_n = \sigma_p$, $\eta_n = \eta_p$, and the expression for θ then reduces to $\eta(\sigma/\lambda)^{\frac{1}{2}}[\frac{1}{2}(T_1 + T_2)]^{\frac{1}{2}}$. For germanium, impurity scattering is dominant in the region of the maximum, so the most favourable thermoelectric power is expected to be about 300 μV. $^{\circ}$C^{-1}. Fig. 5 shows that this prediction is borne out well in practice. This figure also shows that *n*-type germanium is more favourable to thermoelectric refrigeration than *p*-type germanium. It is probable that the main reason for this is a difference in the scattering processes for electrons and holes; for example, in the region where lattice scattering predominates the mobility of electrons in germanium is proportional to $T^{-3/2}$, while the mobility of holes is approximately proportional to $T^{-2\cdot3}$. Such a difference would affect not only conductivity, but also the terms I_n and I_p in the equations for thermoelectric power.

The difference between the properties of *n*-type and *p*-type germanium leads to the important conclusion that, in the

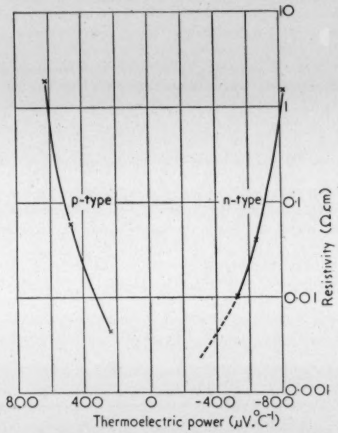


Fig. 4. Thermoelectric power against resistivity for *n*-type and *p*-type germanium. The full curves have been drawn from the experimental results of Middleton and Scanlon[11] and the broken curve has been obtained theoretically

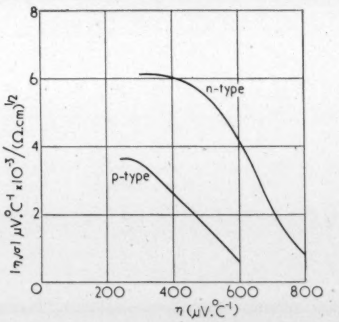


Fig. 5. Variation of the product of thermoelectric power and the square root of the electrical conductivity ($\eta\sqrt{\sigma}$) with the thermoelectric power (η) for *n*-type and *p*-type germanium

search for suitable materials for thermoelectric refrigeration, the *n*-type form of the most favourable semiconductor might well be found to be far superior to the *p*-type form, or vice versa. If this proved to be the case it would be an advantage to make the thermocouples, not out of the *n*- and *p*-type materials, but out of the most favourable form of the semiconductor in conjunction with a metal. Although the metal

would contribute only a negligible part to the total thermoelectric power, the ratio of thermal to electrical conductivity would probably be much less than in the semiconductor. Accordingly the figure of merit, θ, would approach the value of $\eta(\sigma/\lambda)^{\frac{1}{2}}[\frac{1}{2}(T_1 + T_2)]^{\frac{1}{2}}$ for the semiconductor.

For germanium it would certainly be more advantageous to make the thermocouples from the *n*-type material and a metal. It must be pointed out, however, that even then, owing to the high thermal conductivity of germanium (about 0·6 W. cm^{-1}. °C^{-1}) a figure of merit of no more than about 0·13 could be obtained and the maximum cooling would be only about 2° C.

EXPERIMENTAL RESULTS

In view of the criterion previously stated a search was made for semiconducting compound, with a high mean atomic weight, which would be reasonably easy to prepare. One such material is Bi_2Te_3 which can be produced by the direct fusion of the elements. It has been found possible to prepare *p*-type specimens of Bi_2Te_3 with the following properties:

Thermoelectric power = 220 μV. °C^{-1};
Electrical conductivity = 4·0 × 10^2 Ω^{-1}. cm^{-1};
Thermal conductivity = 2·1 × 10^{-2} W. cm^{-1}. °C^{-1}.

If *n*-type specimens of similar properties could be produced the figure of merit for the thermocouple would be 0·51, and a temperature difference of 33° should be attainable. As yet, however, comparable *n*-type material has not been obtained.

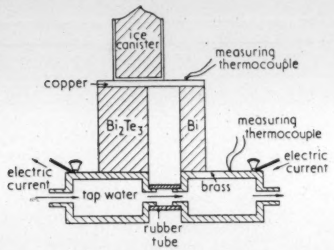


Fig. 6. A refrigerating thermocouple: Bi_2Te_3 and bismuth

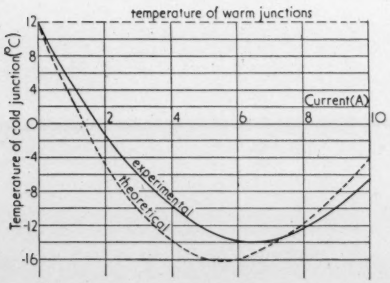


Fig. 7. Variation of cooling from a Bi_2Te_3–Bi thermocouple, with applied current

Dimensions of Bi_2Te_3: 1 × 0·5 × 0·5 cm, and of Bi: 1 × 0·3 × 0·1 cm

Nevertheless a thermocouple has been constructed using the *p*-type Bi_2Te_3 in conjunction with pure bismuth, this couple having a calculated figure of merit of 0·475 and a predicted maximum temperature difference of 28° C. In actual fact a temperature difference of 26° C has been obtained.

The construction of this thermocouple is shown in Fig. 6. Both the elements are 1 cm in length, but whereas the cross-section area of the Bi_2Te_3 specimen is 0·25 cm^2, that of the bismuth specimen is 0·03 cm^2. The maximum temperature difference of 26° C was obtained when a current of about 6·5 A was passed, the thermocouple being lagged with cotton wool. Fig. 7 shows how the temperature difference was found to vary with the current; the theoretical formula,[2] $(T^2 - T_0^2) = (Ri\theta^2/\eta^2)(2\eta T_0 - Ri)$, where R is the resistance of the couple, leads to the broken curve.

In order to demonstrate the refrigerating effect, a small copper canister containing about $\frac{1}{4}$ c.c. of water was placed on top of the copper connecting strip. In about 10 min this water had completely frozen and hoar frost was being deposited on the sides of the container.

CONCLUSIONS

The experiments with Bi_2Te_3 have shown that the thermoelectric refrigeration is now a distinct possibility. It is necessary, however, to increase the maximum temperature difference, at present attainable, so that air-cooling instead of water-cooling can be used. The coefficient of performance, too, is rather low. On the other hand, it must be remembered that there are a large number of compounds of heavy elements which are yet to be investigated. For instance, while the mean atomic weight of Bi_2Te_3 is 160, the mean atomic weights of HgTe and PbTe, both of which are known to be semiconductors, are 164 and 167 respectively. If a ratio of electrical to thermal conductivity four times that for Bi_2Te_3 could be obtained for the same thermoelectric power, the figure of merit, θ, would approach 1·0, and a maximum temperature difference of some 80° C might be possible.

Finally attention is drawn to the fact that the same figure of merit, which applies to thermoelectric refrigerators, also applies to thermoelectric generators, and the latter might well have an important place in the future, notably in connexion with the utilization of solar energy.[13]

REFERENCES

(1) WHITE, W. C. *Elect. Engng, N.Y.*, **70**, p. 589 (1951).
(2) ALTENKIRCH, E. *Phys. Z.*, **12**, p. 920 (1911).
(3) GEHLHOFF, P. O., JUSTI, E., and KOHLER, M. *Abh. Braunsch. Wiss. Gesell*, **2**, p. 149 (1950).
(4) SOMMERFELD, A. *Z. Phys.*, **47**, p. 1 (1928).
(5) EHRENBERG, W. *Proc. Phys. Soc.* [*London*] *A*, **63**, p. 75 (1950).
(6) LAUTZ, G. *Z. Naturforsch.*, **8a**, p. 361 (1953).
(7) FRÖHLICH, H., and MOTT, N. F. *Proc. Roy. Soc. A*, **171**, p. 496 (1939).
(8) CONWELL, E., and WEISSKOPF, V. F. *Phys. Rev.*, **77**, p. 388 (1950).
(9) SHOCKLEY, W., and BARDEEN, J. *Phys. Rev.*, **77**, p. 407 (1950).
(10) GOLDSMID, H. J. *Proc. Phys. Soc.* [*London*] *B*, **67**, p. 360 (1954).
(11) MIDDLETON, A. E., and SCANLON, W. W. *Phys. Rev.*, **92**, p. 219 (1953).
(12) CONWELL, E. M. *Proc. Inst. Radio Engrs*, **40**, p. 1327 (1952).
(13) Report of N.P.L. Committee, *Research*, **5**, p. 522 (1952).

The emission from hot cathodes in gas discharges

By A. E. Pengelly, B.Sc., and D. A. Wright, M.Sc., F.Inst.P., Research Laboratories, The General Electric Co., Ltd., Wembley, Middlesex

[*Paper first received* 7 *April, and in final form* 19 *May*, 1954]

The emission from hot cathodes in gas discharges has been studied over a wide range of current and its stability examined. The condition of zero field thermionic emission has been identified in the presence of a discharge in inert gas. For all types of cathode investigated, the emission at zero field in the discharge has been found similar to that obtained in vacuum. At very high currents, the electron emission is greater than the thermionic emission, either because of enhancement by the field or because of γ-effects at the cathode.

1. Introduction

This paper deals with the electron emission from different types of hot cathode under gas discharge conditions. The cathodes studied include:

(*a*) directly-heated bare filaments of tungsten and thoriated tungsten;
(*b*) directly-heated filaments of tungsten coated with:
  (i) thoria,
  (ii) conventional oxide cathode coating (barium or barium strontium oxide),
  (iii) compounds of barium oxide with other metallic oxides;
(*c*) indirectly-heated nickel cathodes coated with barium or barium strontium oxide.

The oxide coatings were usually prepared by spraying the cathode with the appropriate carbonates, which were decomposed in vacuum. Some cathodes of type (*c*) were studied, however, in which thin films of barium or strontium oxide were evaporated on to the surface of the nickel. The thoria coatings were applied by spraying or by cataphoresis, to a thickness near 0·05 mm. The coatings of type [*b* (iii)] were barium tungstate ($BaWO_4$) and barium zirconate (Ba_2ZrO_4), applied by cataphoresis. The thoriated tungsten filaments were studied both untreated and after carbonizing.

The cathodes for types (*a*) and (*b*) were made of tungsten wire 0·3 mm in diameter, wound to form a helix of diameter 1 mm and length 1 cm. The anode was a nickel plate 2·5 cm square, which could be set parallel with the axis of the helix, spaced 2 to 15 mm away. For type (*c*) the cathodes were formed from a cylindrical nickel box, 3 mm long and 8 mm diameter, the flat top carrying the cathode coating. The box was heated by an alumina-coated molybdenum-tungsten heater mounted inside the box and insulated from it. The anode was parallel with the flat top of the cathode.

The cathodes of type (*c*) were of course unipotential. With types (*a*) and (*b*), the voltage drop across the filament introduced minor difficulties in interpretation. With bare tungsten as type (*a*) at 2500° K this was less than 3 V, however, and with oxide coatings as type (*b*) at 1100° K it was less than 1 V. Comparison of types (*b*) and (*c*) assisted in assessing the importance of this voltage drop on the observed characteristics.

2. Cathode Processing

The diodes were pumped and outgassed in high vacuum. The cathodes were treated in the normal manner; thoria coatings were outgassed at 2000° K, barium tungstate and zirconate at 1500° K, and barium and strontium oxide at 1300° K. The cathodes were given an activating treatment, and d.c. and pulsed emission were measured in vacuum. Spectroscopically-pure inert gas was then introduced as required, and the behaviour under discharge conditions was studied.

3. Temperature Measurement

The cathode temperature was estimated by means of an optical pyrometer as a function of watts with zero anode volts, both in vacuum and after introducing argon at the required pressure. The resistance was also measured with cathodes of types (*a*) and (*b*). With the argon present, and with the cathode at a fixed watts input, the voltage between anode and cathode was increased in stages from zero, in a d.c. circuit with appropriate load resistance. The anode current was plotted against anode voltage, giving curves of the type shown in Fig. 1. The resistance and temperature were observed also as the current increased. Temperatures above about 1400° K could be measured through the argon discharge quite accurately using the optical pyrometer (at 0·665 μ, through the red filter). This was confirmed by working with an a.c. discharge and a stroboscope and comparing the temperatures as observed in the two halves of the cycle, i.e. with and without the glow present. With uncoated filaments agreement was good also with the temperature estimated from the resistance. With the lower temperature cathodes, measurement of temperature in the presence of the discharge was possible only from the resistance for cathodes of type (*b*). For cathodes of type (*c*), thermocouples were welded on the cathode in some of the samples. These were reliable with the evaporated film cathodes, where the emitter and metal temperatures were necessarily similar. With the coatings, however, an appreciable difference is possible between surface temperature and metal temperature.

No conclusions were based exclusively on temperature measurements for coated cathodes of type (*c*).

4. The Form of the Characteristics

Fig. 1 shows small currents *D* limited by electron space charge, before the anode voltage first reached the ionization potential of argon (15·7 V). Near the ionization potential, the current increased rapidly with little increase in voltage. At temperatures lower than the typical operating temperature of each cathode, the current rose in this way to a value such as at one of the points *A* which depended on cathode temperature. At higher temperatures, there was a region of negative slope commencing at the point *B*, followed by a reversal of slope at one of the points *A'*. Beyond the points *A* or *A'*, there was a region over which the logarithm of the current increased linearly with voltage, followed by a region of increasing slope, and then a reversal at *C* or *C'*. Provided that the temperature was high enough to give the first negative slope region, the position of the point *B* and of the plot *BA'* was little influenced by cathode temperature, but the point *A'* moved to higher currents as temperature rose. The suffix 1

refers to temperature T_1, 2 to a higher temperature T_2, and so on.

The interpretation that has been deduced is that at the points such as A and A' the field at the cathode changes from negative (electron retarding) to positive (electron accelerating). The current at such points is therefore that corresponding with zero field at the cathode, where the electron current exceeds the ion current in the ratio $\sqrt{(M/m)}$; M is the mass of the argon atom and m that of the electron. Thus this current should be similar to the zero-field thermionic emission in vacuum. An examination of whether or not this is so is the main feature of this work.

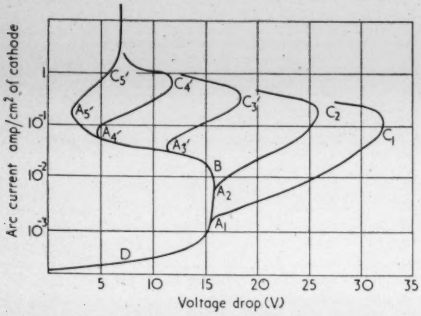
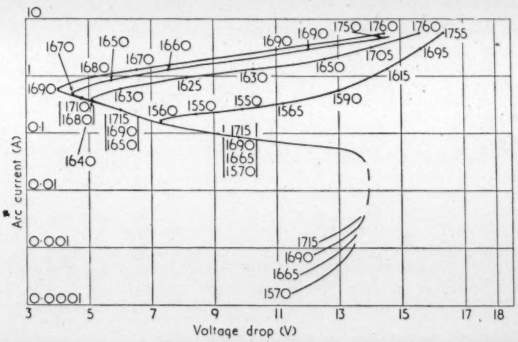


Fig. 1. Diode characteristic for gas discharge tube

Fig. 2. Tube No. 159. Cathode ¼ cm² thoriated tungsten helix (carbonized)

Variation of temperature with arc current. Temperature in °K measured by pyrometer and corrected for emissivity.



Below the points A or A', after once reaching the ionization potential, the current increases rapidly as the electron space charge near the cathode becomes compensated by the positive ion space charge. Beyond the points A or A', further rise with voltage is to be expected because:

(*a*) the field at the cathode is increasing, giving the Schottky decrease in work function;

(*b*) Townsend multiplication processes are occurring, both α-processes in the gas and γ-processes at the cathode.

(*c*) since the ion current and the voltage drop are both rising, the cathode is increasingly heated by ion bombardment. Since the rise in current beyond A or A' is at first exponential with voltage, it appears that here the α-process is the dominating one. The increase in slope at higher current might *a priori* be due either to the γ-processes (*b*) or the heating effect (*c*). This question is discussed further in Section 8.

It was at first difficult to interpret the negative slope region BA', where the voltage drop falls below the first excitation potential of argon (11·4 V). That the cathode field was still retarding was indicated by the fact that in this region, as current rose, the cathode temperature at constant wattage continued to fall, below that at zero current. This was due to the cooling caused by electron liberation. Shortly after passing the point A', the temperature started to rise, no doubt due to ion bombardment. This was studied most fully with thoriated-tungsten filaments, and is illustrated in Fig. 2. It was also noted that at the point B, the visible discharge, which had been located on the anode at lower currents, moved into the cathode anode space, and there took the form of a ball which could be moved about by a magnetic field. It was deduced that there must be a potential maximum at this ball, and that electrons must move from it to the anode under a concentration gradient, against the potential gradient. This reversal of potential gradient would explain the negative slope of the characteristics from B to A' in Fig. 1. Moreover, at B relaxation oscillations set in, of frequency near 2000 c/s, and of amplitude about 4 V. These decreased in amplitude as current increased, become sinusoidal about half-way between B and A', and died out shortly before A' was reached. During the course of this work, Malter, Johnson and Webster[1] published their work on the "ball of fire" mode of glow, confirming these deductions. A theory of the oscillations has been presented recently by Yosimoto.[2] Yosimoto deduced that there are no oscillations when the cathode field is positive. Since the oscillations die out at currents just below A', this confirms the view that A' is the point at which the cathode field is zero.

The region BA' is independent of cathode temperature because the emission is still limited by electron space charge. Here oscillations can occur because the cathode emission can respond at once to any instantaneous change in ion density. The current B, where reversal and oscillations set in, depends on tube and anode geometry, the nature of the gas and its pressure.

Evidently beyond the point A', the cathode field is positive and is increasing with voltage drop in much the same way

as it increases beyond the point A at a lower temperature. The potential maximum must, however, remain present, near to the cathode, since the total voltage drop is at first still less than 11·4 V, the first excitation potential of argon.

5. VACUUM EMISSION

The zero field thermionic emission i_s was measured in vacuum for the various types of cathode, and thermionic constants for the Richardson equation were deduced where the measurements could be made over a sufficient temperature range. This was done in the usual way by plotting $\log i_s/T^2$ against $1/T$ and determining the slope and intercept. No great accuracy is claimed in the present work, especially, of course, for the intercept values. The measurements of work function in a particular tube in a particular state are accurate to $\pm 0{\cdot}15$ eV, and of emission to $\pm 5\%$, but it is often difficult to reproduce the same state at different times or in different tubes. The values obtained are shown in the table, and are emission, and greater therefore than the typical vacuum d.c. emission. With clean tungsten the current remained stable on passing the points such as A, and with thoria coatings there was a short-term stability, though slow changes occurred in accelerating fields as described in Section 7. With all the other cathodes at their lower temperatures, decrease in current occurred rapidly as soon as the field became accelerating. This applied to thoriated tungsten below 1800° K, to oxide coatings below 900° K, and to barium tungstate or zirconate below 1200° K. With thoriated tungsten the deterioration in emission was more marked for the uncarbonized than for the carbonized filaments. The emission was stable with both beyond the points A' at 2000° K, but some decay occurred at 1900° K. With oxide cathodes, the emission fell to about a third of its initial value and was then stable. With the compounds, as with thoriated tungsten below 1800° K, large decreases occurred. Thus bombardment with ions of energy less than 20 V has a marked effect on these cathodes.

The zero-field emission for various types of cathode in vacuum and in argon

Type of cathode	In vacuum				In argon			
	Emission A/cm²	Temp. ° K	Work function	A*	Emission A/cm²	Temp. ° K	Work function	A*
Clean tungsten	0·04	2250	4·4	60	0·04	2250	4·4	60
Thoriated tungsten	2·3	2000	2·6	2	1·9	2000	2·7	3
Thoria coating	2·9	2000	2·6	2·5	0·8 (white)	2000	2·9	4
					1·6 (black)	2000	3·0	14
Sprayed coating								
BaO	0·4	1000	1·4	4·4	0·3	1000	1·2	0·3
BaSrO	3·0	1000	1·1	1·0	0·7	1000	1·1	0·25
Evaporated film								
BaO	0·8	1000	1·2	0·9	0·3	1000	1·1	0·1
BaSrO	1·5	1000	1·2	1·7	2·0	1000	1·1	0·7
SrO on BaO	2·5	1000	1·2	2·8	1·1	1000	1·1	0·4
$BaWO_4$	0·07	1300	—	—	0·3	1300	—	—
Ba_2ZrO_4	0·3	1400	—	—	0·3	1300	1·6	0·3

* Constant in the Richardson equation.

generally similar to the accepted published values (e.g. Wright[3]). The figures for sprayed and evaporated coatings of barium and barium-strontium oxide refer to pulsed emission, which is larger than d.c. at 1000° K. The other figures are for d.c. emission. Following the measurements in high vacuum, gas was introduced, and the same cathodes were studied in the gas discharge.

6. ZERO FIELD EMISSION IN ARGON

If the field at the cathode was not allowed to become positive as the current was raised from zero, the emission was either stable from the start or increased to a stable value. It was possible, therefore, to plot characteristics, stopping near the points A or A' as soon as it was clear that the slope was changing, and so determine the stable zero-field emission i_A at points A or A'. The temperature was also measured there. With these emissions i_A and temperature T, Richardson plots of $\log i_A/T^2$ against $1/T$ were made and the thermionic constants were determined from the plots. Values of the emission and of the constants are shown in the table, and are of the same order in all cases as the values obtained in vacuum. The operation in gas was under d.c. conditions only, and it is noteworthy that, with the barium and barium–strontium oxide coatings, the emission was similar to the vacuum pulsed

7. APPEARANCE OF COATINGS

Thoria coatings were white in appearance after processing and after running with a negative field at the cathode. With accelerating fields and high-current density there was a tendency for the arc to concentrate as field and current increased, leading to overheating of several turns of the filament with respect to the remainder. The overheating was not severely localized, but quite uniform over several turns. When overheating occurred with barium-strontium oxide coatings, however, it was very localized; a few coating particles became intensely heated, and disruption occurred as the current was increased. With the zirconate and tungstate, local heating also occurred, but any particular spot would heat up momentarily, and then cool, a neighbouring spot repeating the process, and so on over several turns of the filament, until after a few seconds the first spot heated up again, and the cycle was repeated. This has been attributed to there being a temperature above which emission decreases, with these coatings, due no doubt to loss of surface barium. There is no such temperature with barium-strontium oxide. With thoria, and with barium tungstate and zirconate, darkening of the coating occurred where overheating was present. The voltage drop rose a little, and the hot region drifted to other parts of the filament so that ultimately the

whole coating was darkened. With thoria, the white state could be restored by exposure to oxygen, either for some hours to air at room temperature, or for a few minutes at 1900° K to an oxygen pressure of 10^{-2} mm. Thus the darkening was due to the formation of excess thorium in the coating. The emission in the dark state was similar to that in the white, frequently a little greater, as shown in the table. With barium–strontium oxide, passage of a high current with the cathode originally at room temperature caused a high temperature cathode spot to move rapidly over the surface of the coating. This caused some disruption, and the coating remaining was darkened and shrunk down on to the metal base. The emission from this coating was similar to that for the original white state.

8. EMISSION IN ACCELERATING FIELDS

As the current was increased beyond the points A or A', under conditions where emission deterioration did not occur, there was at first a range where log (current) increased linearly with voltage drop. Beyond this there was an increase in slope above the linear plot. The whole of this range AC or $A'C'$, where the slope is positive, will be referred to as the first mode of accelerating field operation. Beyond C or C', the slope reverses and the characteristic tends to become a vertical plot, where voltage drop is almost independent of current. This is the second mode of accelerating field operation. At a high temperature such as T_5, this state is established without a reversal, as at C_5, since here the voltage drop is less than that characteristic of the second mode. It was not always possible to trace the full curve; thus only with pure tungsten was there good stability near the point C. With all the other cathodes there were sudden jumps from the first mode near C, and sometimes near C', to the second mode with much lower voltage drop. Moreover, with any of the coated cathodes, operation at high-current density with accelerating field was liable to be interrupted by sudden local overheating, leading to disruption and darkening as discussed. When once local heating had occurred, it became more and more likely to do so in subsequent high current operation.

With clean tungsten, the temperature was followed as arc current was varied between A and C by measuring filament resistance. The increase of current ΔI above the linear logarithmic plot at any point was found to correspond with the increases in temperature ΔT above that at the point A, taking the observed thermionic constants for zero field emission. Thus this part of the curve is due to two processes; α-processes in the gas leading to the linear logarithmic plot, and rise of temperature of the cathode under ion bombardment, causing an increase in zero field emission. The superposition of this effect on that of the α-processes accounts for the curve from A to C; by comparison the effects of γ-processes are not detectable. The behaviour of tungsten in the second mode has been discussed by Druyvesteyn and Penning[4] (p.143). Here the cathode fall is taken to be small compared with the electron free path. This is probably not the case in the first mode.

With thoriated tungsten and thoria coatings in the black state, the same situation as with tungsten seems to apply in the first mode. If the first mode condition between A' and C' was set up, and the input watts to the filament were reduced, maintaining constant arc current, the voltage drop increased and the temperature remained very nearly constant. The increase in watts in the arc compensated almost exactly the decrease in input watts to the filament. Thus in the first mode, the cathode emission is mainly thermionic, and a large proportion of the arc current is carried by ions which reach the cathode.

Operation in the second mode was difficult to achieve with thoriated tungsten and dark thoria coatings. With white thoria it occurred readily, but darkening and a reversion to the first mode, with higher voltage drop, occurred gradually. The barium-oxide type coatings gave only short excursions in the first mode along the curves $A'C'$, and the increase in slope and reversal to the second mode occurred very readily. In the second mode, although the temperature was higher with arc current flowing than without, the temperature alone was by no means sufficient to account for the cathode emission. The intense part of the glow was concentrated near the cathode, giving a very restricted cathode fall region, and an intense field at the cathode. Thus mode two seems to correspond with what is usually described as a field emission arc, though it does not seem clear whether strong field emission or γ-processes are responsible. Currents of the order 10 A/cm^2 have been drawn from barium–strontium oxide cathodes at a temperature of 800° K. Assuming that about a fifth of the current is carried by ions, and applying the space charge laws (e.g. Druyvesteyn and Penning,[4] p. 152), these currents correspond with a dark space charge layer of thickness about 5×10^{-3} cm, and a field at the cathode of about 5×10^5 V/cm. This is not sufficient to produce field emission, but it is sufficient to give a large enhancement of thermionic emission from the coated cathodes, especially barium–strontium oxide, since the increase in emission with field is faster with oxide cathodes than with clean metals.[5] On the other hand, the γ-coefficient for oxide coatings is high,[6] 0·2 for slow positive ions and 0·3 for metastables, so that the γ-process must play some part. It is hoped in future work to obtain further information about its magnitude.

When the watts input to the cathode were lowered when the second mode had been established, there was a fall in temperature. The presence of the arc current maintained the temperature at a fixed amount above that in the absence of arc current, whatever the cathode watts. Moreover, with 7 A arc current, the increase in temperature due to the presence of the arc corresponded with only 16 W of cathode power. Thus in this mode, compared with the first mode, there was either a smaller proportion of the current carried by ions reaching the cathode, or there was a marked reduction in the cathode fall. The latter could occur if the electron emission were due mainly to γ-processes, since these can produce electrons of high initial velocity. If ions in an excited state of energy V_e collide with a surface of work function ϕ, electrons of maximum energy $V_e - \phi$ can be emitted.

9. EFFECTS OF PRESSURE AND COMPOSITION OF THE GAS

Though most of the experiments were carried out using 2 mm of argon, brief studies of other pressures and other gases were carried out. Over the pressure range 1–10 mm with argon, neon, xenon and hydrogen, the zero field emission had similar values to those quoted, and the characteristics were of the same type. The effects on the relationship between cathode watts and temperature were the only ones of significance, though there were small changes in the value of the current at the point B where the first reversal of slope occurred. This is as expected in view of the remarks in Section 3.

10. BARIUM AND STRONTIUM VAPOUR

The presence of vapour of barium and strontium in the discharge was studied spectroscopically with oxide coatings and with barium tungstate and zirconate. With the oxides,

with negative fields at the cathode, barium and strontium lines could be detected at 1100° K, increasing in intensity with temperature. It was necessary, however, to overexpose the argon spectrum by a factor of three to detect the principle lines at 1150° K (barium, 5536 and 4554 Å, and strontium 4607 Å). At any cathode watts, increase of current until the field became positive led to a rapid increase of barium and strontium. For example, with barium-strontium oxide at 1030° K, where the zero field emission was 0·6 A/cm², the barium and strontium lines were as strong as the argon when the current reached 1·4 A/cm², and dominated the discharge with lower cathode watts and similar currents. This, of course, corresponded with incandescent spots in the coatings and a green glow visible to the eye.

With barium tungstate and zirconate, barium was detected with negative fields at the cathode at about 1300° K, and was again increased by operating with a positive field at all temperatures.

11. CONCLUSIONS

The diode characteristics with a hot cathode in a gas discharge have been studied, and the zero-field condition has been identified. The emissions from various types of cathode at zero-field have been found similar to the values obtained in high vacuum.

When the field at the cathode is negative, provided the current density exceeds a certain value depending on tube geometry, gas and pressure, there is a region of negative slope in the diode characteristic, leading to values of voltage drop well below the first excitation potential of the gas. Here there is a potential maximum in the cathode-anode space, and low frequency oscillations usually occur.

When the field at the cathode becomes positive, there is at first a positive slope of the characteristic, where current density rises, because of the combined effects of ionization in the gas and rise of cathode temperature under ion bombardment. At higher currents there is a reversal of slope and finally a variation of current with almost constant voltage drop. Here there is a strong field at the cathode, and the electron emission from it is greater than the thermionic emission, either because of field enhancement of thermionic emission, or because of γ-processes under ion bombardment.

REFERENCES

(1) MALTER, L., JOHNSON, E. O., and WEBSTER, W. M. *RCA Rev.*, **12**, p. 415 (1951).
(2) YOSIMOTO, H. *J. Phys. Soc. Japan*, **8**, p. 59 (1953).
(3) WRIGHT, D. A. *Proc. Instn Elect. Engrs*, **100**, p. 125 (1953).
(4) DRUYVESTEYN, M. J., and PENNING, F. M. *Rev. Mod. Phys.*, **12** (1940).
(5) WRIGHT, D. A., and WOODS, J. *Proc. Phys. Soc.* [*London*] *B*, **65**, p. 134 (1952).
(6) MOLNAR, J. P. *Phys. Rev.*, **83**, p. 940 (1951).

Some properties of symmetrical slit ("cylindrical") electron lenses

By G. D. ARCHARD, A.Inst.P., Associated Electrical Industries Ltd., Aldermaston, Berks.

[*Paper first received* 31 *March, and in final form* 26 *May*, 1954]

The focal lengths of a series of symmetrical slit ("cylindrical") electrostatic electron lenses are determined as a function of aperture size and applied potential. It is shown that all practical lenses may be represented on a single curve by means of a parameter simply related to potential and geometry. From this it is deduced that lens strength depends more on the size of the central aperture than on that of the outer apertures.

The electron optical properties of rotationally symmetrical electron lenses have been widely treated in the literature. Less attention has been paid to the two-dimensional ("cylindrical") lens, possibly because of a relative lack of interest in the production of accurate line foci. Such interest has recently been increasing, for example in connexion with the correction of astigmatism and spherical aberration.[1,2] Exact computational form can be given for the potential distribution of simple cylindrical lenses by means of a conformal transformation; this has been done by Laudet[3] and independently in a slightly different form by the author.[4] For convenience of reference the notation of the latter paper (to be referred to as *I*) will be employed in the deduction of lens properties.

POTENTIAL DISTRIBUTION

The cylindrical lens shown in Fig. 1 has the three equally spaced parallel planes, infinite in extent and indefinitely thin, contain infinitely deep slits set symmetrically about an axis perpendicular to the planes. The outer slits are of breadth $2a$ and the central slit of breadth $2c$. The distance between adjacent planes is b. These dimensions are associated in *I* with certain parameters, m and p, by the relations:

$$\frac{\pi a}{b} = \frac{2(m+p)(1+mp)}{(m^2-1)(p^2-1)} + \log\left[\frac{(m+1)(p+1)}{(m-1)(p-1)}\right] \quad (1)$$

$$\frac{\pi c}{b} = \frac{4(m+p)\sqrt{(mp)}}{(m^2-1)(p^2-1)} + \log\left[\frac{m+p+2\sqrt{(mp)}}{m+p-2\sqrt{(mp)}}\right] \quad (2)$$

The parameters m, p, are then used [*I*, equations (12)–(22)] in the determination of the axial potential and its derivatives. The form of the equations is such that (a/b) and (c/b) may

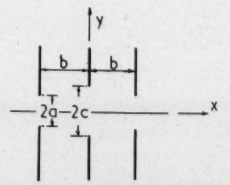


Fig. 1. Cross-section of slit ("cylindrical") lens

easily be determined from m, p, but the reverse process is intractable. It was for this reason that Laudet preferred to plot contours of (a/c), (b/c), in space determined by m, p. Convenient approximations may, however, be developed for the practically useful range of geometries in which (a/b), $(c/b) \leqslant \frac{1}{2}$. Thus, to a first approximation:

$$m = [1 + (\pi/4)^2(c/b)^2]/(\pi/4)(a/b)$$
$$p = m^3/(c/a)^2 \quad (3)$$

For many purposes the formula (3) would prove adequate, as illustrated in Fig. 2; better approximations may always be obtained by resubstituting formula (3) in equations (1) and (2), and using a simple Taylor formula or adjusting by trial. (In all cases worked out in the present paper, adjustment by trial

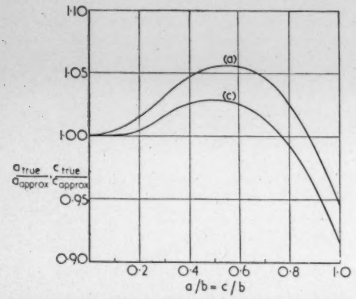


Fig. 2. Relating aperture sizes (a/b), (c/b) given by the true and approximate formulae for the same values of m and p

was continued until the values of m and p, resubstituted in equations (1) and (2), yielded values of a/b, c/b differing from those desired by not more than $\frac{1}{2}\%$.) Owing to the rapid variation of m and p, it will be found convenient to plot the functions $m(a/b)$ and $1/(c/b)\sqrt{(a/b)}\sqrt{p}$ against (a/b) for various values of $(c/b)^2$, which permits nearly linear interpolation. When this has been done, the following information may be extracted: when $a = c$, then $p = m^3$, and, as (a/b) becomes smaller, $(1/m) \propto (a/b)$; when (a/b) is fixed and (c/b) varied, then $(1/p) \propto (c/b)^2$, and m varies only a little; when (c/b) is fixed and (a/b) varied, then $(1/m) \propto (a/b)$ and, for $(a/b) \leqslant \frac{1}{2}$, $(1/p) \propto (a/b)$.

Now expressions (12)–(14) of *I*, which related the axial potential Φ to the axial co-ordinate x by means of a parameter s, may be put into a simpler form by writing:

$$\tfrac{1}{2}\left[\sqrt{(mp)}/s + s/\sqrt{(mp)}\right] = \cosh\theta \qquad (4)$$

Then the potential will become:

$$\Phi = (2V/\pi)\tan^{-1}\left\{\left[2\sqrt{(mp)}/(p-m)\right]\cosh\theta\right\} \qquad (5)$$

where

$$x = \frac{b}{\pi}\left\{\frac{4(m+p)\sqrt{(mp)}}{(m^2-1)(p^2-1)}\sinh\theta + 2\tan^{-1}\left[\frac{2\sqrt{(mp)}}{(p+m)}\sinh\theta\right]\right\} \qquad (6)$$

where the outer electrodes are at potential V and the centre electrode at zero potential. For some applications the outer electrodes may be kept at V_A (accelerating voltage) while the centre electrode is raised above zero (i.e. above cathode potential), so that the potential difference between the electrodes becomes V_A/n. In such a case, equation (5) must be modified so as to reduce $(V - \Phi)$ by the factor n; thus:

$$\Phi = V - (1/n)\left[V - (2V/\pi)\tan^{-1}\left\{\left[2\sqrt{(mp)}/(p-m)\right]\cosh\theta\right\}\right] \qquad (7)$$

Unlike s, θ is a symmetrical parameter; in particular, $\theta = 0$ gives the potential at the lens centre ($x = 0$) as:

$$\Phi_{\text{centre}} = (4V/\pi)\tan^{-1}\sqrt{(m/p)} \qquad (8)$$

With the foregoing empirical relations this shows that, for equal apertures, small compared with their separation, $\Phi_{\text{centre}} \propto (a/b)$. The value of Φ' at the lens centre is plainly zero, while that of Φ'' is:

$$\Phi''_{\text{centre}} = (V/4\pi b^2)(p-m)(m^2-1)^2 (p^2-1)^2/(1+mp)^4\sqrt{(mp)} \qquad (9)$$

For small equal apertures this likewise gives $\Phi''_{\text{centre}} \propto (b/a)$ so that $(\Phi''\Phi)_{\text{centre}} \simeq$ const.

The potential distributions of several lens geometries are variously grouped in Figs. 3–5. In Fig. 3 appear Φ, Φ', Φ''

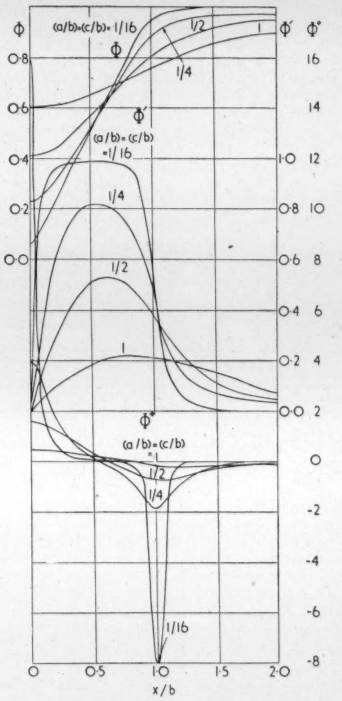


Fig. 3. Potential distribution of cylindrical electron lens $(a/b) = (c/b) = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{16}$

for four equal apertured systems. The proportionality of Φ_{centre} and $1/\Phi''_{\text{centre}}$ to aperture size is seen to be fulfilled more and more closely as (a/b) falls from 1 to $\frac{1}{16}$; at the same time the peak of Φ' moves nearer and nearer to the point midway between centre and outer apertures. In each case the negative peak of Φ'' coincides approximately with the outer aperture and has a value slightly under half that of the positive peak. This helps form a picture of the essential convergence of symmetrical lenses, for it is largely the Φ'' which controls the bending of the electron path.

In Fig. 4, (c/b) is kept constant and (a/b) is varied. The central values of Φ and Φ'' and the peak of Φ' change only

slowly, but the negative peak of Φ'' varies inversely as (a/b). For $(a/b) = \frac{1}{10}$, the negative peak of Φ'' is considerably larger than the positive. The positive area under the curve, however, exceeds the negative, and convergence may still be expected, as will appear.

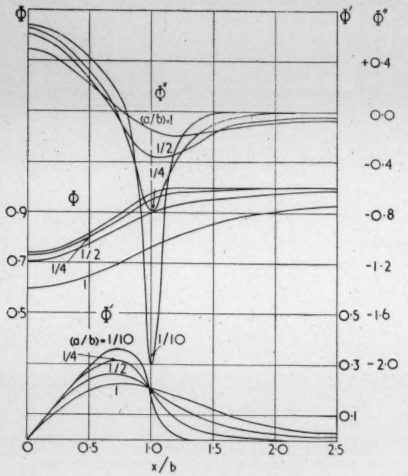


Fig. 4. Potential distribution of cylindrical electron lens $(c/b) = 1$, $(a/b) = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{10}$

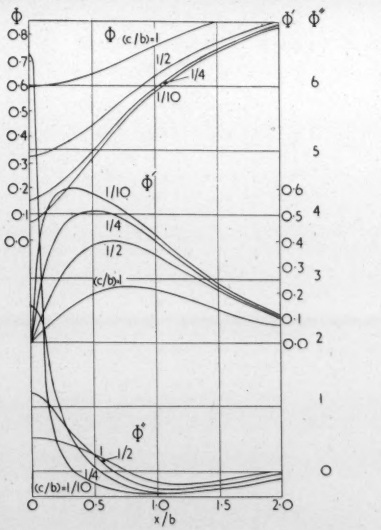


Fig. 5. Potential distribution of cylindrical electron lens $(a/b) = 1$, $(c/b) = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{10}$

In Fig. 5, (a/b) is kept constant and (c/b) is varied. This time the negative peaks of Φ'' change only slowly, while the positive peaks vary inversely as (c/b). The relation $\Phi_{centre} \propto (c/b)$ is obeyed as closely as in the earlier case (Fig. 3) where (a/b) and (c/b) varied together. This suggests that the focal lengths of the lenses will be determined far more by the size of the central aperture than by that of the outer apertures. This will appear later.

FOCAL LENGTHS

The focal lengths of a wide variety of lenses were determined by trajectory plotting (see the Appendix) and are shown in the table. As pointed out in the Appendix, focal length is ordinarily defined with respect to the slope of a ray at the point where it crosses the axis; in some connexions, however, the slope of a ray when it finally leaves the lens is of interest—"focal length" defined with respect to this is distinguished in the table as f'. For weak lenses f and f' are the same.

In the table (a/b) and (c/b) vary from $\frac{1}{10}$ to 1, (c/a) from $\frac{1}{10}$ to 10, and the potential difference V between the electrodes from $0{\cdot}25\ V_A$ to $1{\cdot}55\ V_A$. "Runs" of different values of V/V_A overlap.

It would be particularly convenient if some lens parameter could be found such that all the cases considered could be represented on one graph. Consider the parameter:

$$\eta = \Phi_{[s=\sqrt{(mp)}]}/\Phi_{(s=1)} \qquad (10)$$

s being related to Φ and x by equations (4), (6) and (7). From equation (4), $\Phi_{[s=\sqrt{(mp)}]}$ is clearly the potential at the lens centre, while it is found as a matter of experience that, in all the cases considered, $s = 1$ corresponds very nearly to $x = b$, and that the correspondence becomes closer as aperture sizes diminish. The parameter η is thus approximately equal to the ratio of the potential at the lens centre to that at the centres of the outer electrodes. In terms of m, p and the voltage parameter $n = V_A/V$, η may be written more fully:

$$\eta = \frac{1 - [1 - (4/\pi)\tan^{-1}\sqrt{(m/p)}]/n}{1 - \left\{1 - (2/\pi)\left[\tan^{-1} m + \tan^{-1}(1/p)\right]\right\}/n} \qquad (11)$$

Focal lengths are shown plotted against η in Fig. 6.

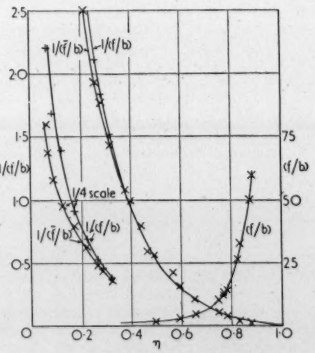


Fig. 6. Focal lengths *versus* parameter η

Despite the wide variations of geometry and potential, all the points lie on a smooth curve, which for approximate purposes may be represented by the empirical formula:

$$1/(f/b) = 0{\cdot}95[(1-\eta)^2 + 3(1-\eta)^3] \qquad (12)$$

within the range $(f/b) = 0{\cdot}7$ to 70 (correct to about 20%).

Focal lengths of cylindrical lenses with various geometries and voltage ratios

(a/b)	(c/b)	V/V_A	$-1/(\bar{f}/b)$	$-1/(f/b)$
$\frac{1}{16}$	$\frac{1}{16}$	1·00	6·39	12·1
$\frac{1}{8}$	$\frac{1}{16}$	1·00	5·51	8·83
$\frac{1}{4}$	$\frac{1}{16}$	0·98	4·67	6·70
$\frac{1}{2}$	$\frac{1}{2}$	1·55	3·82	5·55
$\frac{1}{8}$	$\frac{1}{8}$	1·11	3·15	3·65
$\frac{1}{16}$	$\frac{1}{16}$	0·83	2·51	2·76
$\frac{1}{4}$	$\frac{1}{4}$	1·00	1·93	2·11
$\frac{1}{16}$	$\frac{1}{16}$	0·77	1·77	1·84
$\frac{1}{2}$	$\frac{1}{2}$	1·29	1·43	1·52
$\frac{1}{2}$	$\frac{1}{2}$	1·20	1·08	1·07
$\frac{1}{4}$	$\frac{1}{4}$	0·83	0·984	0·99
$\frac{1}{2}$	$\frac{1}{2}$	1·10	0·792	
1	$\frac{1}{2}$	1·00	0·592	
$\frac{1}{2}$	$\frac{1}{2}$	1·00	0·562	
$\frac{1}{2}$	$\frac{1}{2}$	0·90	0·420	
$\frac{1}{2}$	$\frac{1}{2}$	0·83	0·313	
$\frac{1}{4}$	$\frac{1}{4}$	0·50	0·212	
$\frac{1}{10}$	1	1·00	0·0997	
$\frac{1}{4}$	1	1·00	0·0845	
$\frac{1}{2}$	$\frac{1}{2}$	0·50	0·0834	
$\frac{1}{16}$	$\frac{1}{16}$	0·25	0·0781	
$\frac{1}{2}$	1	1·00	0·0726	
$\frac{1}{4}$	$\frac{1}{4}$	0·25	0·0382	
1	$\frac{1}{10}$	0·25	0·0309	
$\frac{1}{10}$	1	0·50	0·0203	
$\frac{1}{2}$	$\frac{1}{2}$	0·25	0·0168	

This result may be compared with that of Regenstreif[5] for *round* lenses. Using a *simplified* expression for potential distribution, he derived *exact* trajectory equations and plotted the resulting focal lengths against a parameter (x) of the same general nature as η, i.e. (potential at lens centre)/(potential at centre of outer apertures). In the present case *exact* expressions for potential were used and the trajectory equations solved *numerically*; the approximations used in the two cases therefore have some mutual justification. Regenstreif's formula for fairly weak (round) lenses $(x \sim 1)$ was $1/(f/b) = \frac{3}{8}(1-x)^2/x \simeq \frac{3}{8}[(1-x)^2 + (1-x)^3]$, which is of the same general nature as equation (12). The difference in the numerical factors is to be expected as equation (12) relates to cylindrical lenses. It should be mentioned that Regenstreif [his equation (101)] defined focal length with respect to the slope of a ray as it passed the outer aperture. A similar definition in the present case would produce a curve which, for strong lenses, would be below that for $1/(\bar{f}/b)$ in Fig. 6—that is, on the far side of it from the curve for $1/(f/b)$.

It is important for practical purposes to decide whether the results shown in Fig. 6 remain valid when the electrodes have finite thickness. This may be argued by analogy from: (1) Regenstreif, who found that his parameter (x) corresponding to η had a similar form when the central electrode was assigned finite thickness; and (2) MacNaughton[6] who found that the focal length of aperture lenses with finite diaphragm thickness was almost identical with that pertaining to infinitely thin diaphragms.

As an example of the use of Fig. 6, Figs. 7 and 8 have been deduced from it. In determining the values of m, p, and η corresponding to the various geometries and voltages, the original exact equations (1) and (2) were used, so that the

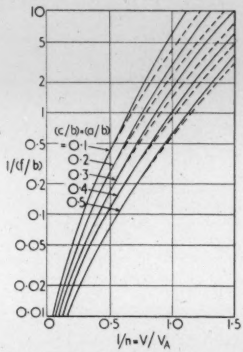


Fig. 7. Focal properties of cylindrical lenses with equal apertures as function of voltage ratio. ——— $1/(f/b)$; – – – – $1/(\bar{f}/b)$

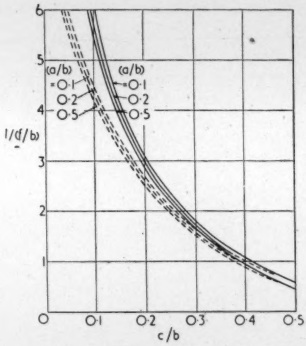


Fig. 8. Focal properties of cylindrical lenses as function of aperture ratio. ——— $1(f/b)$; – – – – $1/(\bar{f}/b)$

only approximation lies in accepting, for the range of interest, the validity of the $\eta - (f/b)$ relation expressed in Fig. 6. In Fig. 7, the apertures a and c are kept equal and the voltage ratio n is varied. In Fig. 8, the voltage ratio is kept constant $(n = 1)$ and (a/b) and (c/b) are varied. This diagram shows clearly that the value of (a/b) plays a relatively small part in the determination of focal length, so that Fig. 7 also would change very little if the restriction $(a = c)$ were removed.

CONCLUSIONS

The variation of focal length with geometry and voltage ratio, presented graphically in Figs. 7 and 8, follows roughly the equation $1/(f/b) = 0{\cdot}95[(1-\eta)^2 + 3(1-\eta)^3]$ where

$\eta \simeq$ (voltage at centre of lens)/(voltage at centre of flank electrodes). b = spacing of electrodes. It is seen that the size of the central aperture ($2c$) controls the refractive power far more than does that of the outer apertures ($2a$). Up to the strongest lens calculated, $1/(f/b)$ increases as (c/b) falls. For stronger lenses, however, the reciprocal of the final slope may be expected to reach a maximum, return to zero (telescopic focusing), and then oscillate, as in the case of Regenstreif.

ACKNOWLEDGEMENTS

The author wishes to thank Mr. M. E. Haine and Dr. G. Liebmann for very helpful comment, Mr. R. K. Taylor for help with some of the computations, and Dr. T. E. Allibone for permission to publish this paper.

REFERENCES

(1) SCHERZER, O. *Optik*, **2**, pp. 114–132 (1947).
(2) SEELIGER, R. *Optik*, **5**, pp. 490–496 (1949); **8**, pp. 311–317 (1951).
(3) LAUDET, M. *Cahiers de Phys.*, **41**, pp. 72–80 (1953).
(4) ARCHARD, G. D. *Brit. J. Appl. Phys.*, **5**, p. 179 (1954).
(5) REGENSTREIF, E. *Ann. Radioélect.*, **6**, pp. 51–83 (1951).
(6) MACNAUGHTON, M. M. *Proc. Phys. Soc.* [*London*] *B*, **65**, pp. 590–596 (1952).

APPENDIX

Plotting of trajectories. Liebmann* has given a procedure for plotting ray paths in round lenses, using a Taylor expansion up to the fourth derivative of the off-axis distance. An analogous procedure may be developed for cylindrical lenses, which require the solution of:

$$\Phi y'' + \tfrac{1}{2}\Phi' y' + \tfrac{1}{2}\Phi'' y = 0 \tag{13}$$

This may be summarized for paraxial rays in the form:

$$y_{n+1} = Q_1 y_n + Q_2 y'_n \tag{14}$$

$$y'_{n+1} = Q_3 y'_n + Q_4 y_n \tag{15}$$

where

$$Q_1 = 1 - \tfrac{1}{4}(\Phi''/\Phi)\Delta x^2 + \tfrac{1}{24}(\Phi'\Phi''/\Phi^2)\Delta x^3 \tag{16}$$

$$Q_2 = \Delta x - \tfrac{1}{4}(\Phi'/\Phi)\Delta x^2 + \tfrac{1}{24}[(\Phi'/\Phi)^2 - 2\Phi''/\Phi]\Delta x^3 \tag{17}$$

$$Q_3 = 1 - \tfrac{1}{4}(\Phi'/\Phi)\Delta x + \tfrac{1}{8}[(\Phi'/\Phi)^2 - 2\Phi''/\Phi]\Delta x^2 \tag{18}$$

$$Q_4 = -\tfrac{1}{4}(\Phi''/\Phi)\Delta x + \tfrac{1}{8}(\Phi'\Phi''/\Phi^2)\Delta x^2 \tag{19}$$

It should be noted that the optical axis of the lens is here assigned the co-ordinate x (instead of the z usual in round lens notation) while the off-axis co-ordinate is called y.

In order to determine focal lengths one should ideally start from a point infinitely remote from the lens, with initial values $y = 1$, $y' = 0$; in practice it is usual to start from a point sufficiently remote from the lens for the potential to be sensibly constant. It is apparent from Figs. 3–5 that Φ' and Φ'' are still appreciable at distances from the lens centre equal to several times the electrode separation. Trial computations, however, show that at distances of the order of $2b$ or $3b$ from the lens centre the $\Phi' y'$ term of equation (13) is very much smaller than the $\Phi'' y$ term; moreover, Φ and y approximately equal unity (taking V as unit potential). Hence, in the outer regions, equation (13) reduces to:

$$y'' + \tfrac{1}{2}\Phi'' = 0 \tag{20}$$

which yields on integration:

$$y' = -\tfrac{1}{2}\Phi' \tag{21}$$

$$y = 1 + \tfrac{1}{2}(1 - \Phi) \tag{22}$$

These may be used as initial values of y, y' in a computation starting at a reasonable distance from the lens (say where $\Phi \geqslant 0{\cdot}9$) and the procedure may be regarded as justified if one or two steps further on equations (21) and (22) still hold. The focal length f is then given by the value of $1/y'$ at the point where $y = 0$.

For some purposes it is convenient to define a different kind of "focal length," namely, the value of $1/y'$ when the electron has passed right through the lens; this is distinguished in the main text as $\bar{f}$. Fig. 6 shows that there is no appreciable difference between f and $\bar{f}$ for lenses having $f \geqslant b$. A further definition, used by Regenstreif [his equation (101)], uses the value of $1/y'$ at the point where $x = b$. This would produce in Fig. 6 a curve $1/(\bar{f}/b)$ a little lower than that given for $1/(\bar{f}/b)$.

A useful check may be derived as follows: equation (13) may be re-written, without any approximation:

$$[\Phi y']_1^2 = -\tfrac{1}{2}[\Phi' y]_1^2 + \int_1^2 (\Phi' y')dx \tag{23}$$

If the lower limit is taken as $x = -\infty$, then $[\Phi y']_1$ and $[\Phi' y]_1$ vanish. If the upper limit is taken as $y = 0$, or as $x = \infty$, then $[\Phi' y]_2$ vanishes. Thus:

$$[\Phi y']_{(2)} = \int_{(x=-\infty)}^{(2)} \Phi' y' dx \tag{24}$$

in which the (2) may be replaced by either ($y = 0$) or ($x = \infty$). Numerical integration of $\Phi' y'$ thus leads to the value of y' either at $y = 0$ (yielding f) or at $x = \infty$ (yielding $\bar{f}$).

For lenses of medium strength ($f \simeq \bar{f}$), equation (24) constitutes the better approach, because $\Phi' y'$ rapidly becomes negligible for $x > 2b$. For very weak lenses, in which y differs little from unity throughout the region of varying potential, the well-known weak lens formula:

$$1/f = \int_{-\infty}^{\infty} (\Phi''/\Phi^{\frac{1}{2}})dx/2\Phi_0^{\frac{1}{2}} \tag{25}$$

may be substituted. [This is obtained by re-writing equation (13) as:

$$\Phi^{\frac{1}{2}}(d/dx)(\Phi^{\frac{1}{2}}dy/dx) = -y\Phi''/2 \tag{26}$$

and setting $y = 1$ on the right-hand side; $\Phi_0 = V_A$ = accelerating potential.]

For very strong lenses it would be expected that the final slope would move back to zero [telescopic focusing: $1/(\bar{f}/b) = 0$] and then oscillate (*cf* Regenstreif[5]). This may be seen as follows: a rough guide to refractive power is $\int \Phi'' dz$ taken over the central aperture; for low (c/b) this approaches Φ'_{peak}. The latter cannot rise above unity (which it would attain for zero aperture), and Fig. 3 shows that even for $a/b = c/b = \frac{1}{16}$ the Φ' curve has started flattening against this upper limit. The turning point had not, however, been reached for the strongest lens calculated ($\eta = 0{\cdot}08$), and in view of the rapidly increasing computational complexity it was not considered expedient to pursue the effect numerically.

* LIEBMANN, G. *Proc. Phys. Soc.* [*London*] *B*, **62**, pp. 753–772, 869 (1949).

A theoretical investigation of the temperature distribution in the metal cutting process

By A. C. Rapier, B.Sc.(Eng.), D.I.C., Plasticity Division, Mechanical Engineering Research Laboratory, East Kilbride, Glasgow

[*Paper first received* 26 *March, and in final form* 18 *May*, 1954]

The problem investigated involves consideration of the temperature distribution in a body moving relative to sources of heat. Relaxation and analytical methods of solution have been developed and applied to the work material, the chip and the tool considered as three separate problems. The results obtained have been compared with other analytical solutions and with indirect experimental evidence of the temperatures attained. The application of the methods has been restricted to the simplest ideal model of the cutting process. The methods may be modified to incorporate the effect of temperature on the thermal properties, yield stress, etc., of the work material and are capable of a wider application to analogous heat flow problems.

1. Introduction

In the metal cutting process the energy dissipated appears largely as heat. The resulting thermal effects are of considerable practical importance. The temperature at the chip-tool interface has a controlling influence on the rate of tool wear, the temperature distribution throughout the work material influences its yield strength and deformation mode, and the temperature cycle through which the material of the new surface has passed determines its metallurgical state.

A precise determination of the temperature distribution by direct experiment, either by optical pyrometers or by thermocouples, has not proved practicable because of the large temperature gradients which exist. These methods give only average values over relatively large areas. The temperature distribution has been determined from indirect evidence, e.g. by inference from metallurgical and other changes in tool and work material. During the last few years, therefore, attention has been directed to theoretical investigations. In the present treatment an attempt has been made to overcome the limitations of the previous solutions which are discussed below.

The idealized model of the metal cutting process considered here is shown in Fig. 1. The workpiece is considered to be moving relative to the cutting tool with a velocity v_1 perpendicular to the edge of the tool so that a layer of material thickness t_1 is removed. The material is assumed to undergo a simple shear on the plane AB and the deformed chip, thickness t_2, width unchanged, then slides up the tool face BC with a velocity v_2; continuity requires that $v_1 t_1 = v_2 t_2$. The energy dissipated is converted into heat by three main mechanisms, namely, by shearing the material along the plane AB, by frictional forces at the tool face BC and subsequently, when a wear band develops from B along the flank of the tool, by friction between the work material and the flank-face. This latter effect has not been considered in previous solutions. Its omission gives the solution for a perfectly sharp tool. It is found that in practice it accounts for only a small proportion of the total energy, and even if it were very much greater, it would have little influence on the shear plane and chip-tool interface temperatures which are of greatest importance. This mechanism of dissipation has therefore been ignored in the present analysis although the methods developed can take it into account if required.

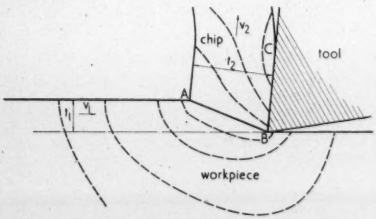


Fig. 1. Idealized model of the metal cutting process
---- = typical isotherms.

Metallurgical examination and deformation patterns on the material show that the deformation is largely confined to two narrow regions, one approximating to the shear plane and the other adjacent to the tool face and that the shear strain rate is uniform across these regions. The shear plane and chip-tool interface are therefore treated as straight line heat sources of uniform strength.

Consider the shear plane source AB. The directions and proportions of the heat-flow from this source depend on the cutting conditions. In general there is a flow of heat into the workpiece, some of which preheats the material to be cut and some heats material below the new surface. Heat is also carried away in the chip and, depending on the temperature gradients, heat may be conducted away from or towards the shear plane inside the chip. On the tool face BC, the extent and distribution of the source is dependent on such factors as the pressure distribution over the contact area and the properties of both tool and work materials at high temperatures. The extent of the wear band observed on tools indicates that the length of BC is of the same order as the chip thickness. The heat produced at this source is conducted into both the chip and the tool. The relative rates of energy input at the two sources can be deduced from the measurements of the forces acting on the tool.

Theoretical solutions for the temperature distribution in the work material, based on Jaeger's[1] solution for a uniform source moving on the surface of a semi-infinite body, have been given by Outwater and Shaw[2] and Loewen and Shaw.[3] The large heat transfer in the direction perpendicular to the shear plane, brought about by the component of the velocity of the material in this direction, makes Jaeger's results inapplicable to this problem. At the lower cutting speeds the temperature distribution, which is obtained along the shear plane, is in error, and at the higher cutting speeds the shear plane temperature and proportion of the heat dissipated, which is carried off in the chip, are both underestimated.

A more reliable solution at the higher cutting speeds has been produced by Hahn[4] by considering the analytical solution for a finite line source moving obliquely through an infinite medium. This infinite medium provides a volume

of material surrounding the source almost twice as great as exists in the actual cutting process. At the lower cutting speeds conduction into the surrounding material accounts for the majority of the heat transferred from the source and the solution is no longer applicable. Hahn's solution also fails to take into account the change of direction of the chip material after it has passed through the shear plane.

Jaeger's solution has been used by Trigger and Chao[5] to obtain the chip-tool interface temperature but only average values were obtained. Since the rate of wear of the tool is governed largely by the maximum temperatures, it is important to obtain the temperature distribution along the interface. The use of the average temperature over the contact area to determine the heat flow into the tool does not give the correct result.

The methods used in the present work enable the temperature distribution throughout the work material and tool to be determined, whereas previously, solutions have only been obtained for the temperatures along the sources. The heating of the material before it is sheared and the heating cycle through which the material of the new surface has passed may be determined from a knowledge of the complete temperature distribution. Further, the methods can be used to take into account such factors as the variation in thermal properties of the work material with temperature, for which only very approximate allowance has been made in previous solutions.

2. THE TEMPERATURE DISTRIBUTION IN A BODY MOVING RELATIVE TO HEAT SOURCES

Consider an infinite body of material moving through fixed heat sources with constant velocity v. Let OX, OY and OZ be rectangular axes fixed relative to the sources, and let the material move in the OX direction. Let the material density be ρ, the specific heat s, the thermal conductivity K, and the thermal diffusivity $\kappa = K/\rho s$.

The differential equation governing the variation of temperature θ with time τ throughout the material is

$$\nabla^2\theta - \frac{v}{\kappa}\frac{\partial\theta}{\partial x} = \frac{1}{\kappa}\frac{\partial\theta}{\partial\tau} \tag{1}$$

If the heat sources are of constant strength, a steady state temperature distribution will be reached which will be determined by the equation

$$\nabla^2\theta - \frac{v}{\kappa}\frac{\partial\theta}{\partial x} = 0 \tag{2}$$

Consider the one-dimensional case of an infinite body of material moving perpendicular to an infinite plane source. Equation (2) then becomes

$$\frac{d^2\theta}{dx^2} - \frac{v}{\kappa}\frac{d\theta}{dx} = 0 \tag{3}$$

for which an exact solution can be obtained. Consider the source, located at $x = 0$, maintained at constant temperature θ_s with sinks maintained at temperature 0 at $x = b$ and $x = -c$.

In the region $x \geqslant 0$

$$\theta/\theta_s = [\exp B - \exp(xB/b)]/(\exp B - 1) \tag{4}$$

where B is the non-dimensional quantity vb/κ.

The rate of heat flow, H_1, per unit area of cross-section in the positive x direction is given by

$$H_1 = \rho v s\,\theta_s - K\frac{d\theta}{dx} = \rho s v\,\theta_s(\exp B)/(\exp B - 1) \tag{5}$$

In the region $x \leqslant 0$

$$\theta/\theta_s = \{\exp[C(x/c + 1)] - 1\}/(\exp C - 1) \tag{6}$$

where $C = vc/\kappa$.

The rate of heat flow, H_2, per unit area of cross-section in the positive x direction is given by

$$H_2 = \rho s v\,\theta_s/(1 - \exp C) \tag{7}$$

Hence the total heat flow from the source per unit area of cross-section is given by

$$H_1 - H_2 = \rho s v\,\theta_s[\exp B/(\exp B - 1) - 1/(\exp C - 1)] \tag{8}$$

Fig. 2 shows the values of θ/θ_s plotted for various values of B and C.

When this solution is compared with the conditions existing near the shear plane in normal metal cutting practice,

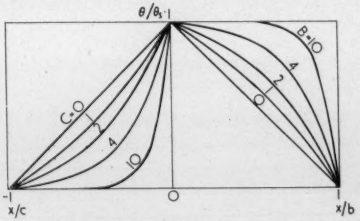


Fig. 2. Temperature distribution in an infinite body moving perpendicular to a plane source

it is found that the values of B and C are of the order 10 to 10^6. The deformed material would, therefore, be raised to the shear plane temperature very rapidly just before reaching the shear plane and would remain at very nearly constant temperature after passing through it, if this were the only source of heat.

The example shows that the direction and magnitude of the temperature gradients do not correspond to the heat flows when the body is moving. When B and C are both infinite H_2 is zero and the total heat flow H_1 from the source is in the positive x direction. If the same temperature distribution existed in stationary material the total heat flow would remain equal to H_1 but would be entirely in the negative x direction. This justifies the statement made above that Jaeger's solution is not applicable to problems in which there is a component of velocity towards the source.

This solution can be generalized to the case of an infinite body of material moving through an infinite plane source inclined at an angle ϕ in the x, y plane to the direction of motion. Consider a system of rectangular co-ordinates x and y with the origin a perpendicular distance u_0 from the source (Fig. 3).

The perpendicular distance of the point (x, y) from the source is

$$u = x\sin\phi + y\cos\phi + u_0 \tag{9}$$

For $u < 0$ and $C \to \infty$

$$\theta/\theta_s = \exp(u\rho s v\sin\phi/K)$$
$$= \exp[v\sin\phi\,(x\sin\phi + y\cos\phi + u_0)/\kappa] \tag{10}$$

since any point in the body can be considered to be moving perpendicular to the source with a velocity $v\sin\phi$. The

component of velocity $v \cos \phi$ parallel to the source does not affect the solution as the isotherms will be parallel to the source and there will be no net heat transfer in this direction.

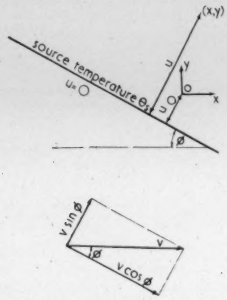


Fig. 3. Notation for a source inclined to the direction of motion of the body

To conform with the notation which has been adopted for this subject, the governing equation may be expressed as

$$\nabla^2\theta - \frac{R}{t}\frac{\partial\theta}{\partial x} = 0 \qquad (11)$$

where $R = vt/\kappa$ and t is the chip thickness.

Relaxation patterns for equation (11) in two dimensions involving the first twelve mesh points (using the standard notation) of a square mesh of side a have been considered. Two of the simplest are shown in Fig. 4 (*a*) and (*b*) together

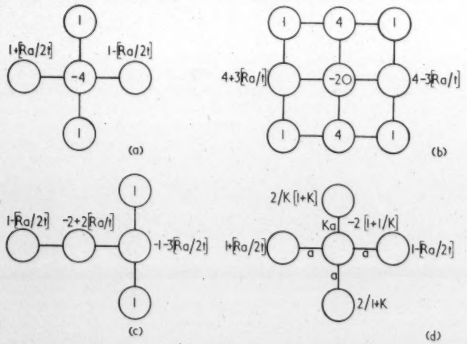


Fig. 4. Relaxation patterns for $\nabla^2\theta - \frac{R}{t}\frac{\partial\theta}{\partial x} = 0$

Error/coefficient of θ_0:

(*a*) $\frac{a^4}{4!}\left(-\frac{1}{2}\frac{\partial^4\theta_0}{\partial x^4} - \frac{\partial^4\theta_0}{\partial y^2\partial x^2} + \frac{1}{2}\frac{\partial^4\theta_0}{\partial y^4}\right)$;

(*b*) $\frac{a^4}{4!}\left(-\frac{3}{5}\frac{\partial^4\theta_0}{\partial x^4} + \frac{3}{5}\frac{\partial^4\theta_0}{\partial y^4}\right)$; (*c*) $\frac{a^3}{3!}6\frac{\partial^3\theta_0}{\partial x^3}$

with the greatest error term divided by the coefficient of θ_0, the value of θ at the point considered. Fig. 4 (*c*) and (*d*) show the modifications to Fig. 4 (*a*) for the cases of a boundary taken arbitrarily in a region where the higher derivatives are known to be small and for a mesh side shortened by a boundary respectively. Experience in the use of these patterns has confirmed that the simplest pattern gives a result to the required accuracy with less labour than more complicated patterns involving up to twelve mesh points.

3. THE TEMPERATURE DISTRIBUTION IN THE WORK MATERIAL

The following simplified model is adopted in the present investigation (see Fig. 5).

(*a*) There is assumed to be no heat loss across the original surface *FA* and the new surface *BD*.

(*b*) It is assumed that *AB* is a line source of strength such that the material is raised to a uniform temperature θ_s on *AB*.

(*c*) The boundary *FE* is assumed to be at the initial workpiece temperature (i.e. zero).

(*d*) The modified pattern [Fig. 4 (*c*)] is used at the boundary *DE*.

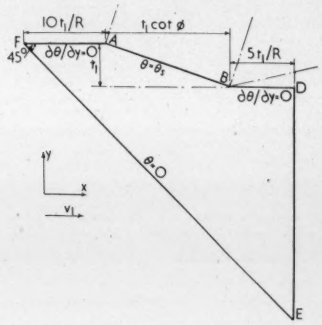


Fig. 5. Boundary conditions for the temperature distribution in the work material

The first assumption is justified for temperatures in the neighbourhood of the cutting edge under all practical cutting conditions. The second assumption is a very good approximation to treating *AB* as a uniform strength source at the higher values of R. If, for the lower values of R, it is required to correct the solution to correspond to a uniform strength source, this correction can easily be applied. Experience of the solutions of this problem over a range of conditions has shown that the indicated positions of the boundaries *FE* and

DE give a good approximation to the actual case of the work material extending infinitely beyond these boundaries. The detailed application of relaxation methods to this problem and calculations of the rates of heat flow into the chip, Q_1, and into the workpiece, Q_2, are not included in this paper.

The temperature increment which would be given to the chip if all the energy used to shear the material was used to heat the chip uniformly would be $\theta_s(Q_1 + Q_2)/Q_1$. This rate of energy input can be deduced from a knowledge of the forces acting on the tool. Having found Q_1/Q_2 for an arbitrary value of θ_s the temperature scale can be adjusted to give the correct value of θ_s.

For the higher values of R it is found that the temperature distribution near the shear plane, except at the extreme ends, approximates closely to the analytical solution for an infinite plane source inclined to the direction of motion, which was obtained above. This enables the heating cycle, through which the material passes before shearing, to be calculated without determining the complete temperature distribution in the work material.

4. THE TEMPERATURE DISTRIBUTION IN THE CHIP

The simplification of the problem which has been investigated is shown in Fig. 6.

The boundary conditions are taken as follows.

(*a*) The material leaves the shear plane AB at a uniform temperature θ_s.

(*b*) There is a uniform heat source BC along the tool face for a distance αt_2

$$\text{i.e.} \quad -\partial\theta/\partial y = R\theta_f/\alpha t_2$$

where θ_f is the temperature increment which would be given to the chip by this source if the chip were heated uniformly. The frictional work done at the chip-tool interface can be

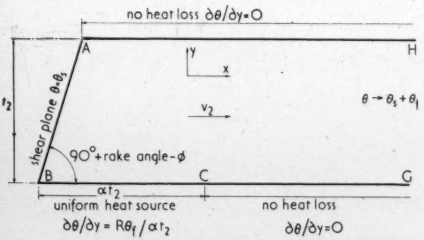


Fig. 6. Boundary conditions for the temperature distribution in the chip

determined from a knowledge of the forces acting on the tool. θ_f is calculated by assuming, as will be shown later to be approximately true, that all this work is used to heat the chip.

(*c*) There is no heat loss from the remaining chip surfaces AH and CG, i.e. $\partial\theta/\partial y = 0$. This can be justified over the region in which we are interested.

(*d*) As $x \to \infty$ the temperature across the chip becomes uniform, i.e. $\theta \to \theta_s + \theta_f$. By taking a boundary a distance $2\alpha t_2$ from the shear plane the modified pattern [Fig. 4 (*c*)] can be used with sufficient accuracy.

As R increases the value of $\partial\theta/\partial x$ becomes relatively small over most of the region and conduction in the x direction can then be neglected. An analytical solution can then be determined by considering a strip in the y direction which moves over the source located at $y = 0$ and is insulated from the material on either side.

The problem is then to solve $\dfrac{\partial^2\theta}{\partial y^2} = \dfrac{\rho s}{K}\dfrac{\partial\theta}{\partial\tau} = \dfrac{R}{v_2 t_2}\dfrac{\partial\theta}{\partial\tau}$ for the conditions

$$\theta = \theta_s \text{ for all } y \text{ at } \tau = 0$$

$$\frac{\partial\theta}{\partial y} = -\frac{R\theta_f}{\alpha t_2} \text{ at } y = 0 \text{ from } \tau = 0 \text{ to } \tau = \frac{\alpha t_2}{v_2}$$

$$\frac{\partial\theta}{\partial y} = 0 \begin{cases} \text{at } y = 0 \text{ from } \tau = \dfrac{\alpha t_2}{v_2} \text{ to } \tau = \infty \\ \text{at } y = t_2 \text{ from } \tau = 0 \text{ to } \tau = \infty. \end{cases}$$

Making the substitution $x = \tau v_2$ the solution for $\tau = 0$ to $\tau = \alpha t_2/v_2$ can be expressed as:

$$\frac{\alpha}{R}\frac{\theta - \theta_s}{\theta_f} = \frac{(t_2 - y_2)^2}{2t_2^2} + \frac{x}{Rt_2} - \frac{1}{6} + \frac{2}{\pi^2}\sum_{n=1}^{n=\infty} \exp(-n^2\pi^2 x/Rt_2)\cos(yn\pi/t_2)/n^2 \quad (12)$$

This converges very rapidly for the values of R in which we are interested.

In most cases only the chip tool interface temperature distribution is of interest, i.e. the values for $y = 0$. This can be expressed approximately as

$$\frac{\alpha}{R}\frac{\theta - \theta_s}{\theta_f} = 1{\cdot}13\sqrt{\left(\frac{x}{Rt_2}\right)} \quad (13)$$

5. THE TEMPERATURE DISTRIBUTION IN THE TOOL

The problem to be solved is shown in Fig. 7. Since the material is stationary relative to the heat sources, the governing equation for the steady state temperature distribution becomes

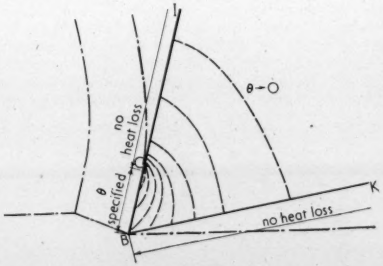


Fig. 7. Boundary conditions for the temperature distribution in the tool

- - - - = typical isotherms.

simply $\nabla^2\theta = 0$. The temperature distribution along the contact surface with the chip BC is taken from the solution for the chip. Heat losses from the tool surfaces CI and BK are neglected. A satisfactory approximation to the remaining boundary condition, that the temperature should tend to

the ambient temperature at a large distance from the cutting edge, can be obtained by choosing a boundary sufficiently far from the cutting edge.

Solutions have been obtained for particular cases and show that in general the temperature gradients perpendicular to the contact surface *BC* are very much greater on the chip side than on the tool side. This, together with the fact that the work material usually has a higher thermal conductivity than the tool material, means that the majority of heat generated flows into the chip.

6. APPLICATION OF METHODS TO PARTICULAR PROBLEMS

(*a*) *The temperature distribution in the work material.* Fig. 8 shows the relaxation solution for the conditions $R = 1$, $\tan\phi = \frac{1}{3}$ for a uniform shear plane temperature of 1000

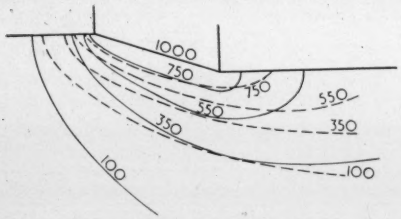


Fig. 8. The temperature distribution in the work material

——— = relaxation solution $Q_1/(Q_1 + Q_2) = 0{\cdot}373$;
- - - - - = solution by Chao and Bisacre.

arbitrary units compared with the numerical solution by Chao and Bisacre.[6] The discrepancy is probably due to the greater degree of approximation of their method. Relaxation solutions for other values of R and ϕ show that the proportion of the heat produced at the shear plane which is carried away in the chip, $Q_1/(Q_1 + Q_2)$ increases with increase of R and with increase of ϕ.

The proportion of heat carried away in the chip, A, according to Hahn[4] can be deduced from Fig. 5 of his paper by taking the dimensionless average temperature of the shear plane divided by the dimensionless adiabatic temperature. (The values of the dimensionless adiabatic temperature are in error and should be multiplied by a factor

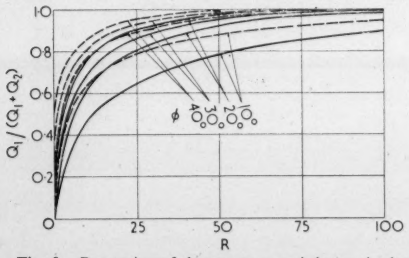


Fig. 9. Proportion of shear energy carried away in the chip

——— = adapted from Hahn's results;
- - - - - = corrected values.

$\cos^2\phi$ giving $\pi/\tan\phi$.) On the assumption that in Hahn's solution an equal proportion of the total heat generated is conducted away perpendicular to the direction of motion on each side of a line corresponding to *FABD* in Fig. 5, it can be shown that $Q_1/(Q_1 + Q_2) = 2A/(A + 1)$. This assumption is approximately true for all values of ϕ and R. Hahn's original values and the corrected values are shown in Fig. 9. Chao and Trigger[7] have attempted to correct Hahn's solution by multiplying the source strength by a factor equivalent to $2Q_2/(Q_1 + Q_2)$. While Hahn's solution is a good approximation for the higher values of R, their solution does not give an accurate result under any conditions.

The following table is a comparison of the values of $Q_1/(Q_1 + Q_2)$ obtained by relaxation methods with the values given by Loewen and Shaw and the values given by Hahn.

ϕ	18·4°	18·4°	9·5°
R	1	2	1
Relaxation	0·373	0·501	0·302
Loewen and Shaw	0·303	0·380	0·235
Hahn (corrected)	0·212 (0·350)	0·319 (0·483)	0·186 (0·314)

There is good agreement between the relaxation value and the corrected Hahn solution. The corrected curves of Fig. 9 can therefore be used for the approximate determination of shear plane temperatures in practical problems where the temperature distribution throughout the work material is not of importance and the full relaxation solution is not required.

(*b*) *The temperature distribution in the chip.* A valuable piece of indirect evidence of the temperature distribution along the chip-tool interface is given by Trent.[8] He studied the wear of cemented carbide tools and produced evidence to show that appreciable wear only occurred where the interface temperature exceeded 1300° C. Calculation of the temperatures for the conditions of Trent's tests by the methods given above show that:

(1) temperatures exceeding 1300° C will be produced over almost exactly the proportion of the contact area on which cratering was observed by Trent;
(2) the temperature gradients into the tool are very much smaller than the gradients into the workpiece for this range of cutting conditions. The majority of the heat produced at the interface goes into the chip which justifies the statement made in Section 5 above;
(3) as a consequence of this, the thermal conductivity of the tool has, for the higher values of R, little control on the temperatures at the interface and the increases in tool material conductivity which could be made in practice would not appreciably decrease tool wear as suggested by Trent. Also the temperature distribution inside the tool and any accompanying softening effects will not vary appreciably with tool conductivity;
(4) a further consequence of the small proportion of the heat going into the tool is that steady conditions at the chip-tool interface are reached very rapidly, probably after only a few times the contact length of chip has crossed the contact face, i.e. in about 10^{-3}s at 300 ft/min in this instance.

Calculations on these lines give a quantitative explanation of effects of speed and feed for which Trent was unable to give an account. Trent's work provides satisfactory confirmation of the theoretical approach used.

It is possible, using the methods suggested, to obtain a more satisfactory estimation of the temperatures involved

in metal cutting than has been given by the approximate analytical methods previously applied to this problem. The boundary conditions of the simplified model, line sources and uniform thermal properties assumed here are not implicit in the methods used. The complete temperature distribution can be obtained for any experimental data. This will enable the influence of temperature on all aspects of machinability, particularly tool wear, to be determined.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance of Mr. W. S. McDougall and Mr. T. R. Cook in some of the computation work. The work described was carried out in the Plasticity Division of the Mechanical Engineering Research Laboratory as part of the research programme of the Mechanical Engineering Research Board. It is published by permission of the Director of the Laboratory.

REFERENCES

(1) JAEGER, J. C. *Proc. Roy. Soc. N.S.W.*, **76**, p. 203 (1942).
(2) OUTWATER, J. O., and SHAW, M. C. *Trans Amer. Soc. Mech. Engrs*, **74**, p. 73 (1952).
(3) LOEWEN, E. G., and SHAW, M. C. *Trans. Amer. Soc. Mech. Engrs*, **76**, p. 217 (1954).
(4) HAHN, R. S. *Proc. U.S. National Congress of Applied Mechanics*, p. 661 (New York: American Society of Mechanical Engineers, 1951).
(5) TRIGGER, K. J., and CHAO, B. T. *Trans Amer. Soc. Mech. Engrs*, **73**, p. 57 (1951).
(6) CHAO, B. T., and BISACRE, G. H. *Proc. Instn Mech. Engrs*, **165**, p. 1 (1951).
(7) CHAO, B. T., and TRIGGER, K. J. *Trans Amer. Soc. Mech. Engrs*, **75**, p. 109 (1953).
(8) TRENT, E. M. *Proc. Instn Mech. Engrs*, **166**, p. 64 (1952).

The two-dimensional electric field of a curved-sided wall or groove on an infinite plane

By N. H. LANGTON, Ph.D., A.Inst.P., and N. DAVY, D.Sc., The University, Nottingham

[*Paper first received* 18 *December*, 1953, *and in final form* 29 *January*, 1954]

The electric or magnetic field of a two-dimensional system consisting of a curved-sided wall or groove on an infinite plane is investigated theoretically using conformal transformations and elliptic integrals. Graphs and tables showing the variation of field intensity are given, and tables of the elliptic integrals K, E, K' and E' for complex modulus k, where $k^2 = 1 + \exp(ix)$ are appended.

In previous papers on the subject of two-dimensional fields,[1,2,3] the authors have used the method applicable to curvi-linear triangles involving the use of the integral solution of the hypergeometric series. In the problems described here, which again involve a curvi-linear triangle, a different method of solution is used. It is shown here how the solution of a new problem can be obtained from the solution of a previous problem by the use of complex geometrical inversion. In this way, much of the difficulty in obtaining the initial geometrical transformation is avoided.

The variation of the electrical intensity R along the axis of symmetry OG of the system is given, as well as that along the edge of the conductor, because the variation along OG is that most likely to be of practical importance. As in the papers referred to above, the calculation requires values of the elliptic integrals for complex moduli. Since such values are not easily obtainable, tables of E,K,K' and E' for complex moduli are appended.

THE TWO-DIMENSIONAL FIELD OF A CURVED-SIDED WALL OR GROOVE

The systems, the two-dimensional fields of which are to be obtained, are shown in Figs. 2 and 5. Both diagrams are drawn in the z_1-plane, and represent cross-sections of thick conductors with a surface charge, or slabs of magnetized ferromagnetic material in an unsaturated condition, with either a curved projection or groove running along the length. The sides of the projection or groove BOE are quarter circles, and DB and ED are straight lines proceeding to infinity at D. In both diagrams the unshaded area is that which contains the field. The systems could also be very thin charged or magnetized plates. The electric fields only will be investigated here, although the same mathematics is equally applicable to the magnetic case. With slight modifications the theory can also be applied to the flow of incompressible liquids or gases past the systems, or to the flow of heat from the bodies.

The first part of the solution consists in a transformation from the z_1-plane to a c-plane, where $c = k^2$ and k is the elliptic modulus, so that the edge of the conductor becomes

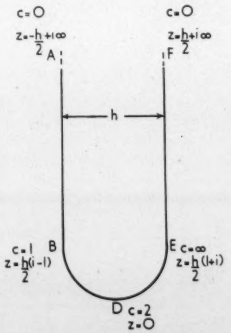


Fig. 1.
The z-plane

the real axis of the c-plane and the area containing the field corresponds to the upper half of this plane. This geometrical transformation could be performed in the manner described in a previous paper,[1] as the diagram in the z_1-plane is a curvilinear triangle. It is more convenient, however, to obtain it indirectly. It is easy to verify that the line $DBOED$ of Fig. 2 or Fig. 5 can be obtained by an inversion of the

U-shaped figure, Fig. 1, about the point D, according to the equation

$$z.z_1 = h^2/4 \qquad (1)$$

Since this inversion is complex, it also involves a reflexion in the real z-axis, giving the figures required. The area external to the U-shape of Fig. 1 is transformed into the unshaded area of Fig. 5, and the internal area of the U-shape becomes the unshaded area of Fig. 2. If the distance between

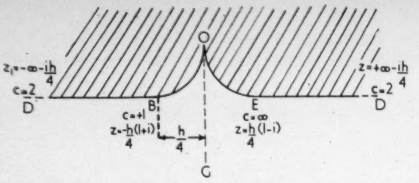


Fig. 2. The z_1-plane

the parallel straight sides AB and FE is h, then the radius of the arcs OB and OE is $h/4$. Hence both the geometrical transformations can be obtained.

THE FIELD OF THE CONDUCTOR OF FIG. 2

The equation that transforms the boundary and the internal area of the conductor of Fig. 1 into the real axis and upper half of the c-plane has been shown to be

$$z = ih.(K' - E')/E + h(i - 1)/2 \qquad (2)$$

where the z-origin is at the point D. In this equation, K, K', E and E' are complete elliptic integrals of the first and second kinds, defined as

$$K = \int_0^1 \frac{dx}{\sqrt{[(1-x^2)(1-k^2x^2)]}}$$

$$K' = \int_0^1 \frac{dx}{\sqrt{[(1-x^2)(1-k'^2x^2)]}}$$

$$E = \int_0^1 \frac{\sqrt{(1-k^2x^2)}dx}{\sqrt{(1-x^2)}} \qquad E' = \int_0^1 \frac{\sqrt{(1-k'^2x^2)}dx}{\sqrt{(1-x^2)}}$$

where k is real and $0 < k < 1$, and $k^2 + k'^2 = 1$.

Combining equation (1) with equation (2) gives, as the required transformation equation

$$z_1 = h/[4i(K' - E')/E + 2(i - 1)] \qquad (3)$$

The real c-axis, with the appropriate values of c marked, is shown in Fig. 3. From the distribution of the corner points

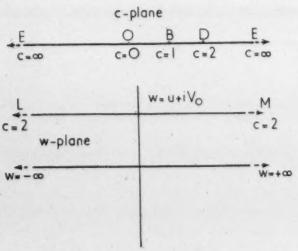


Fig. 3. The c- and W-planes

it can be seen that the unshaded portion of Fig. 2, which contains the field we are to investigate, corresponds to the upper half of the c-plane.

Equation (3) can be checked by verifying that it gives the correct values of z_1 at the corners.

When $c = 0$, $(K' - E')/E = \infty$ and hence $z_1 = 0$, which is correct.

When $c = 1$, $(K' - E')/E = 0$ and hence $z_1 = -h(1 + i)/4$, which is correct.

When $c = \infty$ we must use the values of K', E' and E for modulus k greater than unity. When we do this we obtain $z_1 = h(1 - i)/4$, which is correct.

When $c = 2$ we similarly obtain $z_1 = \infty$, which is the correct value for the corners D.

The values of z_1 calculated from equation (3) are given in the first two lines of Tables 1 and 2. These points, when plotted, give the correct outline OBD.

The electrical intensity R. In order to obtain the expression for R, it is necessary to effect a second transformation to link the electric properties of the field with the geometric properties of the conductor. Let the potential of the edge $DBOED$ be $+V_0$ e.s.u. everywhere. The potential V of all points outside the conductor will be less than $+V_0$ and range downwards to a value of $-\infty$ e.s.u. The diagram of Fig. 3 shows the W-plane, or electric plane, where $W = U + i.V$, U being a stream function and V a potential. The horizontal line LM has an equation $W = U + i.V_0$ and extends to infinity at each end, which corresponds to $c = 2$. If we can set up a one-to-one correspondence between the W-plane and the c-plane so that the line LM becomes the real axis of the c-plane and the area below LM corresponds to the upper half of this plane, then the required electrical conditions will be obtained. This correspondence is given by the transformation

$$dW/dc = A/(c - 2)^2 \qquad (4)$$

where A is a constant.

The electrical intensity R is given by $R = |dW/dz_1|$.

Now $dW/dz_1 = (dW/dz)(dz/dz_1)$ and $dz/dz_1 = -h^2/4z_1^2$

Also $dW/dz = (dW/dc)(dc/dz)$ and $dc/dz = -4cE^2/ih\pi$

Hence $$R = (hA/\pi)|cE^2/(c - 2)^2.z_1^2| \qquad (5)$$

in terms of z_1, or, in terms of z,

$$R = (16A/\pi h^3)|cE^2z^2/(c - 2)^2| \qquad (6)$$

This value of R may be tested for correctness at the corners.

At the corner O, $z_1 = 0$, $c = 0$, which gives $R = 0$, which is correct.

At the corner B, $z_1 = -h(1 + i)/4$, $c = 1$, and hence $R = 8A/\pi h$.

At the corner E, $z_1 = h(i - 1)/4$, $c = \infty$ Here we must use the value of E for modulus greater than unity. It is more convenient to change to modulus c' using the relation $c = c' + 2$ which means we have to evaluate the limit

$$(\pi/hA).R = \underset{c' \to \infty}{Lt} |(E^2/c').(c' + 2)/c'z_1^2|$$

It can be shown that $\underset{c' \to \infty}{Lt} E^2/c' = -1$, and when $c' = \infty$ we know that $z_1 = -h(1 + i)/4$ so we obtain $R = 8A/\pi h$, which is the same as the value at the corner B.

To obtain the value of R at the corner D where $c = 2$ we must use a similar procedure, using equation (6), when we obtain the value $R = 3{\cdot}438(8A/\pi h)$.

The tables below show the values of R, calculated from equation (6), along the curved side OB, the straight side BD and along the centre line OG.

Calculation of R along the curved side OB. Along *OB* the values of c vary from $c = 0$ to $c = +1$ and hence we can use equation (6) as it stands. R is tabulated in terms of the constant $(\pi h/16A)$ and z_1 in terms of h. These results are plotted on the graph of Fig. 5 the abscissae being the real part of z_1.

Table 1 (see also Fig. 4). *The variation of R along OB*

c	=	0·0000	0·0076	0·0301	0·0669
z_1/h	=	0·0000	−0·0226 } −0·1037i }	−0·0338 } −0·1255i }	−0·0448 } −0·1462i }
$R(\pi h/16A)$	=	0·000	0·026	0·0695	0·1196
c	=	0·2500	0·5000	0·5878	0·7500
z_1/h	=	−0·0802 } −0·1832i }	−0·129 } −0·219i }	−0·1402 } −0·2244i }	−0·1765 } −0·2388i }
$R(\pi h/16A)$	=	0·2742	0·409	0·443	0·498
c	=	0·8830	0·9046	0·9700	1·000
z_1/h	=	−0·2110 } −0·2470i }	−0·2172 } −0·2480i }	−0·2310 } −0·2498i }	−0·2500 } −0·2500i }
$R(\pi h/16A)$	=	0·526	0·5275	0·520	0·500

The graph shows that R increases from zero to a maximum value a short distance from the corner B.

Calculation of R along the straight side DB. Along *BD* the values of c vary from $+1$ at B to $+2$ at the point D at infinity. The results of the calculations are shown in Table 2. The curve resulting from these values is shown in Fig. 4, where R is in terms of the constant $(\pi h/16A)$ and the abscissae are the real parts of z_1 in terms of h.

Table 2 (see also Fig. 4). *The variation of R along BD*

c	=	1·0000	1·030	1·137	1·334
z_1/h	=	−0·2500 } −0·250i }	−0·264 } −0·250i }	−0·314 } −0·250i }	−0·470 } −0·250i }
$R(\pi h/16A)$	=	0·500	0·482	0·451	0·437
c	=	1·490	1·600	1·703	2·000
z_1/h	=	−0·673 } −0·250i }	−0·918 } −0·250i }	−1·331 } −0·250i }	∞ } −0·250i }
$R(\pi h/16A)$	=	0·436	0·432	0·416	1·719

The values show that R continues to fall as we move away from the corner B for a considerable distance, and then rises to the value of $1{\cdot}719(16A/\pi h)$ at infinity. Only the first part of the curve can be shown on the graph. This type of variation is what would be expected, since the charge will move away from the sharp corner O and be drawn towards the corners B and E.

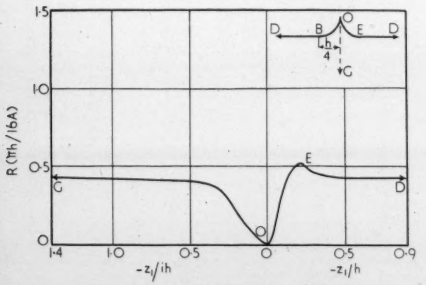


Fig. 4. Variation of R along *DEOG*

Calculation of R along the centre line OG. To investigate the variation of R along this line, we notice that it has the c-values of 0 and $+2$ at its ends. Thus *OG* corresponds to some line joining the points O and D in the c-plane, and it can be proved that this line is a semi-circle, centre at the point B on the real c-axis, and radius unity. The equation of this semi-circle is $c = 1 + \exp(\mathrm{i}x)$, where $\pi \geqslant x \geqslant 0$. Thus the points on the line *OG* are given by using complex values of c given by the above expression. The values of the elliptic integrals for various values of x are given in the Appendix. Table 3 gives values of z_1 and R for complex c.

Table 3 (see also Fig. 4). *Variation of R along OG*

c	=	2·000	1·9877 } +0·1564i }	1·890 } +0·458i }	1·588 } +0·810i }
z_1/h	=	$-i\cdot\infty$	−3·289i	−1·208i	−0·692i
$R(\pi h/16A)$	=	1·719	0·395	0·421	0·414
c	=	1·309 } +0·951i }	1·000 } +1·000i }	0·5458 } +0·890i }	0·293 } +0·707i }
z_1/h	=	−0·554i	−0·441i	−0·363i	−0·305i
$R(\pi h/16A)$	=	0·410	0·398	0·375	0·350
c	=	0·0488 } +0·309i }	0·000 } +0·000i }		
z_1/h	=	−0·220i	−0·000i		
$R(\pi h/16A)$	=	0·254	0·000		

The curve illustrating this variation of R along the centre line *OG* shows that R rises from zero value at the corner O to reach a maximum at a distance of about h from O, after which there is a slow decrease in value to the point G at infinity. This is the type of variation that would be expected.

The value of the constant A in the expression $R(\pi h/16A)$ can be found as follows. If q is the charge on a strip of conductor, the depth normal of which to the z_1-plane is unity,

then $$q = (W_2 - W_1)/4\pi$$

where W_2 and W_1 are the W-values at the ends of the strip. Now $W = A/(c-2) + i.V_0$, where V_0 e.s.u. is the potential of the conductor. At the origin O we have $c = 0$, hence $W_1 = -A/2 + i.V_0$. At the corner B, $c = 1$ so that $W_2 = -A/2 + i.V_0$. Hence $W_2 - W_1 = A/2$ and therefore $A = 8\pi q$.

THE FIELD OF THE CONDUCTOR OF FIG. 5

To obtain the geometrical transformation which makes the edge *DBOED* become the real axis of the c-plane and the area above the conductor (unshaded) become the upper half of the c-plane, we use the inversion equation $z.z_1 = h^2/4$,

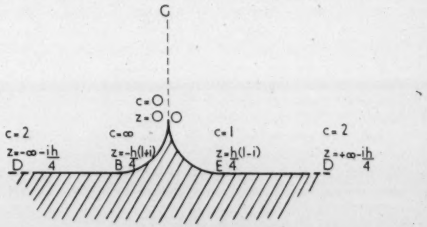


Fig. 5. The z_1-plane

where $$z = ih.\beta/\alpha + h(1+i)/2 \qquad (7)$$

and $$\alpha = (2-c)E - 2c'.K, \quad \beta = (2-c)E' - c.K'$$

It has been shown[2] that equation (7) transforms the U-shaped figure so that the area external to it becomes the upper half of the c-plane. Therefore the equation

$$z_1 = h/[4i\beta/\alpha + 2(1 + i)] \qquad (8)$$

will transform the line *DBOED* into the real axis of the c-plane so that the unshaded area corresponds to the upper half of the c-plane. The origin of the co-ordinates is the point O in Fig. 5, and the point D in Fig. 1. The only difference between this inversion and the one for the previous problem, is that in this case the values of c at the corners of the U-shaped conductor are slightly different. In this case, $c = \pm\infty$ at the corner B, and $c = +1$ at the corner E, thus making the external area and not the internal one correspond to the upper half of the c-plane. A similar exchange of values is necessary along the real c-axis, shown in Fig. 3, to make it correspond to this problem. The geometrical transformation given by equation (8) can be checked by verifying that it gives the correct corner values.

It is easy to verify that the values of z_1, when $c = 0$, $c = +1$ and $c = 2$, are correct. If $c = \infty$, $K = K' = 0$, $E = i.\infty$ and $E' = \infty$. To find the value of z_1 we must evaluate the limit $Lt_{c\to\infty}\ \beta/\alpha$. This limit can be found by first putting $c = 1/c_1$ in the expression for β/α and rearranging, when we obtain $\beta/\alpha = [E' - K'/(2c_1 - 1)]/(E - 2K)$. When $c = \infty$ or $c_1 = 0$ then $K = K' = 0$ and in effect we have to evaluate the limit $Lt_{c_1\to 0}\ E'/E$ which is the same as $Lt_{c\to\infty}\ E'/[E - c_1K - i(E' - c_1K')]$, using the formulae for E and E' valid for when the modulus is greater than unity. Hence we obtain $Lt_{c_1\to 0}\ E'/E = +i$. Therefore, when $c = \infty$ we obtain $z_1 = -h(1 + i)/4$ which is the correct value for the corner B. Table 4 shows the values of z_1 calculated from equation (8).

The electrical intensity R. We use the same electrical transformation as in the previous problem [equation (4)]. In this case, however, equation (7) gives

$$dz/dk = -3ck\pi ih/2\alpha^2$$

Hence

$$R = (Ah/3\pi).|\alpha^2/z_1^2.c.(c - 2)^2| \qquad (9)$$

where A is a constant, and z_1 is given by equation (8).

At the corner E where $c = +1$, $z_1 = -h(1 + i)/4$ and $\alpha = +1$. Hence $R = 8A/3\pi h$.

At the corner B where $c = +\infty$, $z_1 = -h(1 + i)/4$ and $\alpha = \infty$.

To find the value of R we must evaluate the limit of $\alpha^2/z_1^2.c.(c - 2)^2$ as c approaches infinity. To do this we must rewrite this expression in the form which is valid for modulus k greater than unity and then find the limit as $c \to 0$. When this is done we obtain $R = 8A/3\pi h$.

Hence the two values of R at the corners E and B are the same, so that equation (9) appears to be satisfactory.

At the corner O, $c = 0$ and $z_1 = 0$. We must now find the limit of equation (9) when c approaches zero. When α and z_1 are written in full, we find without any difficulty that $R = \infty$ at the corner O as would be expected.

At the corner D where $c = 2$ we must again use the modified equation and find the limit. This procedure gives the value $R.(3\pi h/A) = 5{\cdot}741$. Tables 4, 5 and 6 show the values of R calculated from equation (9) for other points along the edge and also along the centre line.

Calculation of R along the curved side OE. Along OE c varies from $c = 0$ to $c = +1$ and hence the calculation of R from equation (9) is straightforward. The results are shown in Table 4, where R is tabulated in terms of the constant $3\pi h/A$. To show the results graphically $R(3\pi h/A)$ is plotted vertically against the real values of z_1 in Fig. 6.

Table 4 (see also Fig. 6). *Variation of R along OE*

c	=	0·0000	0·410	0·5000	0·7500
z_1/h	=	0·0000	0·00001	0·0026	0·0356
		−0·0000i	−0·0203i	−0·0363i	−0·128i
$R(3\pi h/A)$	=	∞	27·82	19·93	9·14

c	=	0·883	0·970	1·000
z_1/h	=	0·123	0·2152	0·2500
		−0·215i	−0·2478i	−0·2500i
$R(3\pi h/A)$	=	6·502	7·17	8·00

These results show that R drops from an infinite value at the corner O to a minimum value just before the corner E is reached. The curve then rises until the corner E is reached. The next set of results show that the value of R continues to rise for some distance beyond this corner.

Calculation of R along the straight side ED. Along ED the values of c vary from $c = +1$ to $c = 2$. Thus we must use a modified form of equation (9) obtained by expressing the elliptic integrals in the forms valid for moduli greater than unity. Table 5 shows the results of these calculations, which are graphed in Fig. 6 . As before, $R(3\pi h/A)$ is plotted against the real part of z_1.

Table 5 (see also Fig. 6). *Variation of R along ED*

c	=	1·000	1·030	1·137	1·334
z_1/h	=	0·250	0·285	0·406	0·678
		−0·250i	−0·250i	−0·250i	−0·250i
$R(3\pi h/A)$	=	8·000	9·01	10·72	11·98

c	=	1·702	2·000
z_1/h	=	2·092	∞
		−0·250i	−0·250i
$R(3\pi h/A)$	=	12·51	5·741

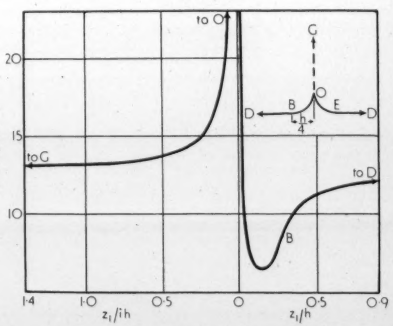


Fig. 6. Variation of R along *DBOG*

Calculation of R along the centre line OG. The procedure is the same as that described for the previous problem. Again we make the transformation $c = 2/(1 + c_2)$ and use values of the elliptic integrals for complex modulus of the form $c = 1 + \exp(ix)$.

These results are shown in Table 6, and the results plotted in the usual way on Fig. 6.

Table 6 (see also Fig. 6). *Variation of R along OG*

c	= 2·000	1·9877 } +0·1564i }	1·951 } +0·3090i }	1·890 } +0·4580i }
z_1/ih	≐ ∞	4·44	2·06	1·27
$R(3\pi h/A)$	= 5·741	12·7	13·0	13·2
c	= 1·707 +0·5878i }	1·707 +0·707i }	1·588 +0·810i }	1·309 +0·951i }
z_1/ih	= 0·90	0·665	0·511	0·285
$R(3\pi h/A)$	= 13·3	13·4	13·8	14·5
c	= 1·000 +1·000i }	0·000 +0·000i }		
z_1/ih	= 0·195	0·000		
$R(3\pi h/A)$	= 15·7	∞		

The curve obtained from these results, labelled *OG* in Fig. 6, shows that R drops quickly from its infinite value at the sharp corner O as we move along *OG*. At a distance of about h from O, the intensity reaches an almost steady value, and from here drops very slowly as we move towards G.

COMPARISON WITH LEES' WALL PROBLEM

The electrostatic field of a vertical conducting wall of finite height standing on an infinite conducting plane was obtained by Lees[4] for work in connexion with the potential gradients of the earth's atmosphere. It is interesting to compare the results of the last problem (Fig. 5) with the results of Lees. The system consists of a vertical wall of height h standing on a conducting plane, and a cross-section of the conductor is shown in Fig. 7, where *BOE* is the wall and *DBED* repre-

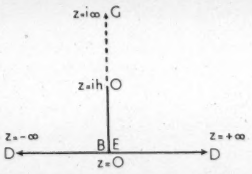


Fig. 7. The vertical wall, z-plane

sents the plane. The system *DBOED* can be regarded as one plate of a condenser, the other plate being at infinity. Lees shows that the line *DBOED* is transformed into the real axis of a c-plane by the equation

$$z = h\sqrt{(c^2 - \alpha^2)}/\alpha$$

where α is a constant. Lees takes the system to be at zero potential, whilst we shall make the potential $+V_0$ e.s.u. as in the previous problems. The electrical transformation is therefore

$$W = -A.c + iV_0$$

In the usual way we now find the electrical intensity R to be given by

$$R = (h/A\alpha).|z|/\sqrt{(z^2 + h^2)} \qquad (10)$$

From this equation the values of R along the sides *DE*, *EO* and up the continuation of the wall *OG* can be obtained. Since these values are easy to obtain, they are not given here. Values of $R(h/A\alpha)$ from the above equation have been obtained, together with values of z, measured from the origin of the system, which is the point E (or B), and there results are shown plotted in Fig. 8, together with the curve of Fig. 6 for comparison. This latter curve is shown dotted. To render comparison easier, the curve of Fig. 6 has been altered in scale. This is done as follows.

Equation (10) shows that, at a great distance from the wall, $(z = \infty)$, $R(h/A\alpha) = 1{\cdot}000$. Table 5 shows that, in the case of the curved wall, when $z = \infty$, we have $R(3\pi h/A) = 5{\cdot}741$. Thus by dividing all the values of $R(3\pi h/A)$ by 5·741 in Tables 4, 5 and 6 the intensities at infinity will be the same. This is equivalent to adjusting the value of the constant A

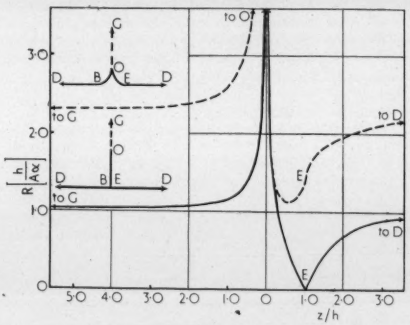


Fig. 8. Variation of R for straight and curved wall

in the expression for $R(3\pi h/A)$, or to adjusting the surface density of charge. Since the height of the curved wall is $h/4$, the z-values in Tables 4, 5 and 6 have been multiplied by 4 before plotting.

Fig. 8 shows that, under these conditions, the intensity is always greater on the curved wall than at an equivalent point on the straight wall system. In practice, it would be very difficult to make the join at the base of the wall a perfect right angle, and the slightest radius at this join would cause the intensity variation to tend to follow the dotted curve. In practice, the curve of intensity variation would lie somewhere between the extremes of Fig. 8.

REFERENCES

(1) DAVY, N., and LANGTON, N. H. *Brit. J. Appl. Phys.*, **3**, pp. 156–8 (1952).
(2) DAVY, N., and LANGTON, N. H. *Quart. J. Mech. Appl. Math.*, **6**, Pt. 1, pp. 115–21 (1953).
(3) LANGTON, N. H., and DAVY, N. *Brit. J. Appl. Phys.*, **4**, pp. 134–7 (1953).
(4) LEES, C. H. *Proc. Roy. Soc.*, **91**, pp. 440–51 (1915).

APPENDIX

In this Appendix some values of the elliptic integrals K, E, K' and E' are given for complex modulus k. Some of these values were recently calculated by Cambi,* upon whose results these tables are based. Table 7 gives values of complex moduli k of the form $k^2 = c = 1 + \exp(ix)$, where x is measured in right angles. Cambi does not tabulate values of E or E', and also gives K and K' in the form $r.\exp(i\theta)$,

* CAMBI, E. *J. Math. Phys.*, **26**, No. 4, p. 234 (1948).

whilst for purposes in connexion with the type of calculation met with in two-dimensional field problems, it is more convenient to have the values in the form $(a + ib)$ as given below. The magnitude and square of the integrals, which are frequently required, are also given below. The values are given to four figures. Since the calculations were performed using four figure tables or a twenty-inch slide rule, the last figure is doubtful. The accuracy, however, is sufficient for most purposes in connexion with electric or magnetic fields.

Cambi gives K for $k^2 = 1 + \exp(ix)$ and K' for $k^2 = 1 + \exp(ix)$. From these values Tables 8 and 10 are obtained. To calculate the values of E and E', use is made of Cambi's tables giving $\exp(iy/2).[(K' - E')/K']$ for $k^2 = -\exp(ix)$. Cambi gives corresponding values of x and y, and since K' is known, we can obtain values of E' for $k'^2 = 1 + \exp(ix)$. These values are the same as those for E for $k^2 = 1 + \exp(ix)$. Hence Table 9 is constructed.

Values of E for $k'^2 = 1 + \exp(ix)$ can now be calculated using Legendre's relation

$$E'.K + E.K - K.K' = \pi/2$$

Hence Table 11 is obtained.

Table 7. *Values of x, c and c'*

x in right angles	$c = 1 + exp\,(ix)$	$c' = 1 - c = -exp\,(ix)$
0·0	2·000	−1·0000
0·1	1·9877 + 0·1564*i*	−0·9877 − 0·1564*i*
0·2	1·9511 + 0·3090*i*	−0·9511 − 0·3090*i*
0·3	1·8910 + 0·4540*i*	−0·8910 − 0·4540*i*
0·4	1·8090 + 0·5878*i*	−0·8090 − 0·5878*i*
0·5	1·7071 + 0·7071*i*	−0·7071 − 0·7071*i*
0·6	1·5878 + 0·8090*i*	−0·5878 − 0·8090*i*
0·7	1·4540 + 0·8910*i*	−0·4540 − 0·8910*i*
0·8	1·3090 + 0·9511*i*	−0·3090 − 0·9511*i*
0·9	1·1564 + 0·9877*i*	−0·1564 − 0·9877*i*
1·0	1·0000 + 1·0000*i*	− 1·0000*i*
1·1	0·8436 + 0·9877*i*	+0·1564 − 0·9877*i*
1·2	0·6910 + 0·9511*i*	+0·3090 − 0·9511*i*
1·3	0·5460 + 0·8910*i*	+0·4540 − 0·8910*i*
1·4	0·4122 + 0·8090*i*	+0·5878 − 0·8090*i*
1·5	0·2929 + 0·7071*i*	+0·7071 − 0·7071*i*
1·6	0·1910 + 0·5878*i*	+0·8090 − 0·5878*i*
1·7	0·1090 + 0·4540*i*	+0·8910 − 0·4540*i*
1·8	0·0489 + 0·3090*i*	+0·9511 − 0·3090*i*
1·9	0·0123 + 0·1564*i*	+0·9877 − 0·1564*i*
2·0	0·0000 + 0·0000*i*	+1·0000

Table 8. *Values of K, K^2, $|K|$ for $c = 1 + e^{ix}$, or of K', K'^2 and $|K'|$ for $c' = 1 + exp\,(ix)$*

x in right angles	*K or K'*	*K^2 or K'^2*	*$|K|$ or $|K'|$*
0·0	1·311 + 1·311*i*	3·438*i*	1·854
0·1	1·338 + 1·236*i*	0·236 + 3·310*i*	1·821
0·2	1·362 + 1·164*i*	0·503 + 3·175*i*	1·792
0·3	1·386 + 1·092*i*	0·728 + 3·060*i*	1·765
0·4	1·409 + 1·026*i*	0·933 + 2·854*i*	1·744
0·5	1·428 + 0·955*i*	1·127 + 2·730*i*	1·718
0·6	1·447 + 0·8874*i*	1·306 + 2·568*i*	1·698
0·7	1·467 + 0·822*i*	1·477 + 2·412*i*	1·681
0·8	1·482 + 0·755*i*	1·626 + 2·240*i*	1·663
0·9	1·495 + 0·690*i*	1·763 + 2·062*i*	1·646
1·0	1·508 + 0·6255*i*	1·882 + 1·886*i*	1·631
1·1	1·522 + 0·561*i*	1·995 + 1·704*i*	1·622
1·2	1·532 + 0·4962*i*	2·098 + 1·524*i*	1·610
1·3	1·541 + 0·4345*i*	2·186 + 1·338*i*	1·601
1·4	1·550 + 0·3701*i*	2·276 + 1·147*i*	1·591
1·5	1·556 + 0·3094*i*	2·326 + 0·962*i*	1·587
1·6	1·562 + 0·2475*i*	2·379 + 0·774*i*	1·581
1·7	1·565 + 0·1850*i*	2·415 + 0·578*i*	1·576
1·8	1·568 + 0·1203*i*	2·444 + 0·377*i*	1·573
1·9	1·569 + 0·0617*i*	2·458 + 0·198*i*	1·571
2·0	$\pi/2 = 1·5708$	$\pi^2/4 = 2·468$	1·5708

Table 9. *Values of E, E^2, $|E|$ for $c = 1 + exp\,(ix)$, or of E', E'^2 and $|E'|$ for $c' = 1 + exp\,(ix)$*

x in right angles	*E or E'*	*E^2 or E'^2*	*$|E|$ or $|E'|$*
0·0	0·5991 − 0·5991*i*	− 0·7178*i*	0·8473
0·1	0·6745 − 0·6241*i*	0·0655 − 0·8420*i*	0·9190
0·2	0·7470 − 0·6430*i*	0·1440 − 0·9620*i*	0·9856
0·3	0·8206 − 0·6498*i*	0·2512 − 1·0660*i*	1·047
0·4	0·8960 − 0·6515*i*	0·3783 − 1·169*i*	1·108
0·5	0·9675 − 0·6457*i*	0·5235 − 1·248*i*	1·163
0·6	1·035 − 0·6320*i*	0·6800 − 1·310*i*	1·212
0·7	1·104 − 0·6180*i*	0·8370 − 1·363*i*	1·265
0·8	1·168 − 0·5946*i*	1·005 − 1·389*i*	1·311
0·9	1·228 − 0·5652*i*	1·189 − 1·388*i*	1·352
1·0	1·284 − 0·5320*i*	1·364 − 1·364*i*	1·389
1·1	1·339 − 0·4945*i*	1·532 − 1·318*i*	1·429
1·2	1·383 − 0·4477*i*	1·712 − 1·238*i*	1·454
1·3	1·4261 − 0·4020*i*	1·871 − 1·147*i*	1·482
1·4	1·464 + 0·3537*i*	1·997 − 1·036*i*	1·504
1·5	1·4959 − 0·2987*i*	2·1484 − 0·8908*i*	1·525
1·6	1·524 − 0·2417*i*	2·264 − 0·7406*i*	1·543
1·7	1·542 − 0·1828*i*	2·346 − 0·5638*i*	1·552
1·8	1·557 − 0·1260*i*	2·409 − 0·3928*i*	1·562
1·9	1·567 − 0·0614*i*	2·452 − 0·1925*i*	1·568
2·0	$\pi/2 = 1·5708$	$\pi^2/4 = 2·468$	1·5708

Table 10. *Values of K, K^2, $|K|$ for $c = -exp\,(ix)$, or of K', K'^2 and $|K'|$ for $c' = -exp\,(ix)$*

x in right angles	*K or K'*	*K^2 or K'^2*	*$|K|$ or $|K'|$*
0·0	1·311	1·719	1·311
0·1	1·313 − 0·0276*i*	1·723 − 0·0725*i*	1·313
0·2	1·315 − 0·0558*i*	1·726 − 0·1467*i*	1·316
0·3	1·321 − 0·0828*i*	1·738 − 0·2188*i*	1·324
0·4	1·327 − 0·1132*i*	1·748 − 0·3008*i*	1·332
0·5	1·337 − 0·1412*i*	1·769 − 0·3780*i*	1·345
0·6	1·346 − 0·1710*i*	1·783 − 0·4602*i*	1·356
0·7	1·360 − 0·2007*i*	1·807 − 0·5459*i*	1·376
0·8	1·376 − 0·2312*i*	1·839 − 0·6370*i*	1·395
0·9	1·397 − 0·2627*i*	1·882 − 0·7340*i*	1·421
1·0	1·421 − 0·2950*i*	1·932 − 0·8390*i*	1·451
1·1	1·453 − 0·3292*i*	2·004 − 0·9570*i*	1·490
1·2	1·485 − 0·3640*i*	2·071 − 1·080*i*	1·528
1·3	1·528 − 0·4008*i*	2·174 − 1·224*i*	1·580
1·4	1·580 − 0·4403*i*	2·302 − 1·392*i*	1·640
1·5	1·645 − 0·4810*i*	2·475 − 1·597*i*	1·714
1·6	1·731 − 0·5257*i*	2·720 − 1·818*i*	1·809
1·7	1·850 − 0·5760*i*	3·101 − 2·132*i*	1·937
1·8	2·025 − 0·6300*i*	3·703 − 2·562*i*	2·121
1·9	2·342 − 0·6950*i*	5·002 − 3·257*i*	2·443
2·0	∞	∞	∞

Table 11. *Values of E, E^2, $|E|$ for $c = -exp\,(ix)$ or of E', E'^2 and $|E'|$ for $c' = -exp\,(ix)$*

x in right angles	*E or E'*	*E^2 or E'^2*	*$|E|$ or $|E'|$*
0·0	1·9104	3·648	1·9104
0·1	1·906 + 0·0460*i*	3·631 + 0·1754*i*	1·909
0·2	1·897 + 0·0932*i*	3·589 + 0·3533*i*	1·899
0·3	1·884 + 0·1372*i*	3·530 + 0·5173*i*	1·889
0·4	1·867 + 0·1816*i*	3·453 + 0·6780*i*	1·875
0·5	1·839 + 0·2225*i*	3·332 + 0·8190*i*	1·853
0·6	1·806 + 0·2600*i*	3·194 + 0·9302*i*	1·825
0·7	1·772 + 0·2915*i*	3·005 + 1·033*i*	1·795
0·8	1·730 + 0·3240*i*	2·888 + 1·121*i*	1·766
0·9	1·684 + 0·3495*i*	2·713 + 1·172*i*	1·720
1·0	1·633 + 0·3692*i*	2·531 + 1·206*i*	1·674
1·1	1·578 + 0·3800*i*	2·346 + 1·194*i*	1·623
1·2	1·521 + 0·3872*i*	2·163 + 1·178*i*	1·569
1·3	1·461 + 0·3900*i*	1·983 + 1·139*i*	1·512
1·4	1·401 + 0·3860*i*	1·814 + 1·082*i*	1·454
1·5	1·333 + 0·3715*i*	1·644 + 0·9720*i*	1·384
1·6	1·265 + 0·3283*i*	1·492 + 0·8310*i*	1·307
1·7	1·200 + 0·2750*i*	1·368 + 0·6600*i*	1·233
1·8	1·134 + 0·1871*i*	1·251 + 0·4242*i*	1·149
1·9	1·067 + 0·0990*i*	1·129 + 0·2112*i*	1·072
2·0	1·0000	1·0000	1·0000

Correspondence

The use of carbon crucibles in measurements on the rate of evaporation of liquid metals in vacuum

I have read with interest the paper by M. G. Rossmann and J. Yarwood published in the January issue of this *Journal*. In the discussion on bubble formation in heated liquids it is suggested that some workers have neglected the possible presence of bubbles in calculating rates of evaporation. Further, that they have calculated the evaporation rates of liquid metals from their vapour pressures; presumably this is intended to mean calculations based on the assumption of free evaporation from a surface.

The treatment given to the subject by Rossmann and Yarwood is based entirely on establishing the hydrostatic pressures at different depths in the liquid metal and does not relate this to the temperature distribution required in the liquid metal for the postulated bubble to, in fact, be formed. If this had been done it would have been obvious that the high thermal conductivity of a liquid metal does not at normal rates of evaporation allow for bubble formation.

Consider, for example, liquid aluminium with a surface temperature of 810° C at which the vapour pressure is of the order 10^{-4} mm of mercury.(1) Then for a bubble to be formed with a vapour pressure of 10^{-3} mm of mercury, corresponding to a temperature of 890° C, a hydrostatic pressure of 6×10^{-3} mm of aluminium would be necessary (the density of liquid aluminium in this temperature region has been reported(2) as 2·3 g/cm^3). Thus a temperature gradient between the evaporating surface and the site of bubble formation would need to be $80° C/6 \times 10^{-3}$ mm (i.e. 133 000°C/cm). The thermal conductivity of aluminium at 800° C is 0·21 cal per sec per cm^2. Thus with the temperature gradient postulated above heat energy would be supplied to the evaporation surface at a rate of 26·6 kcal per sec per cm^2.

At a surface temperature of 810° C the evaporation rate of aluminium is 10^{-6} g per cm^2 per sec, which, ignoring heat losses by radiation, would require a heat energy input of 3×10^{-3} cal per cm^2 per sec to supply the latent heat of evaporation. Thus in the foregoing example bubble formation would not be possible. Inspection of the vapour pressure-temperature values and thermal conductivities of metals will show that the foregoing is not a special case but a general characteristic of metal evaporation.

Bubble formation may only occur where heat energy is supplied at an *abnormally* high rate and the conditions of heating give rise to an uneven temperature distribution in the liquid metal. Disruption of a liquid metal surface is more likely to arise from expansion of absorbed gases in the metal. It is also possible that bubble formation may occur at the surface of a liquid metal which is covered by an impervious layer of low volatility, e.g. surface oxide.

Research Laboratory, L. HOLLAND
W. Edwards and Co. (London) Ltd.,
Crawley, Sussex.

REFERENCES

(1) DUSHMAN, S. *Scientific Foundations of Vacuum Technique*, 1st Ed., p. 748. (London: Chapman and Hall, Ltd., 1949.)
(2) *Liquid Metals Handbook*. (Washington: Atomic Energy Commission, Dept. of the Navy, 1950.)

Though one must agree with Mr. L. Holland's interesting calculations regarding the formation of bubbles in a heated metallic liquid in a vacuum, yet it is not considered that he has disproved the simple account put forward by M. G. Rossmann and myself. Holland, in effect, postulates that a temperature gradient must exist in the liquid metal before a bubble can form. We assumed that a bubble could form below the liquid surface, even if the metal were at a uniform temperature throughout. Thus, suppose the temperature of, for example, liquid aluminium is 996° C so that its vapour pressure is 10^{-2} mm of mercury,* and that the pressure of the gas above the aluminium is negligible in comparison, i.e. it is in a high vacuum. Then, assuming that the density of aluminium is 2·0 gm cm^{-3}, this would imply that the aluminium *could* exist in the vapour state at depths below the surface up to $10^{-2} \times 13{\cdot}6/2 = 6{\cdot}8 \times 10^{-2}$ mm.

If some slight inhomogeneity were to exist at such a depth, then there is no reason why a bubble should not form there and expand as it rises to the surface. Hence the surface cannot be considered as plane. The real question is: what are the physical conditions which lead to bubble formation in a hot liquid metal? We have assumed that they are definitely produced at small depths in a liquid at uniform temperature. There seems to be no reason why Holland should suppose that they are not.

Department of Mathematics and Physics, J. YARWOOD
The Polytechnic,
Regent Street,
London, W.1.

The Editor has been good enough to let me see a proof of Mr. Yarwood's reply and I wish to comment as follows: in my discussion of bubble formation in liquid metals it was assumed that the pressure at a given depth in the liquid was solely the hydrostatic pressure h. In actual fact there will also be a pressure on the surface of the liquid due to the vapour atoms. If the vapour and liquid were in *equilibrium* the pressure on the liquid surface would be equal to the saturated vapour pressure P_v. Under conditions of free evaporation in a vacuum, where the vapour emission is not impeded by either gas or vapour atoms, the pressure due to the vapour would be equal to $0{\cdot}5\,P_v$, because vapour atoms impart an impulse to the surface in the act of escaping but do not as in *equilibrium* evaporation impinge on the liquid surface. Thus the pressure acting in depth in the liquid is $h + 0{\cdot}5\,P_v$. It can be seen that allowance for this additional pressure term will make bubble formation all the more unlikely. This treatment also ignores the effect of surface tension forces against which a bubble would have to expand.

In my example the vapour pressure of the postulated bubble was very much greater than the reaction pressure of the vapour atoms and the latter could safely be neglected. This is not so in Yarwood's example and its neglect has lead

* Using Dushman's figures [*Scientific Foundations of Vacuum Technique*, p. 748 (London: Chapman and Hall, Ltd., 1949)] for want of better information. Actually aluminium in the liquid state is an unsatisfactory metal on which to base an argument on account of its marked ability to wet all surfaces in which it comes into contact. However, this does not concern the point at issue here and, in any case, silver, gold and mercury were discussed in the original paper.

to a fallacious result. The true pressure acting at a depth of $6{\cdot}8 \times 10^{-2}$ mm in the liquid aluminium would be $1{\cdot}5 \times 10^{-2}$ and not 10^{-2} mm of mercury as stated by Yarwood. An approximate calculation shows that for bubble formation at this depth the temperature would need to be at least 7° C higher than that of the surface, which is extremely unlikely to occur.

L. HOLLAND

This correspondence is now closed.—Ed.

Note on the resistance-network analogue solution of field problems of spherical symmetry

It was shown in a preceding paper* that the design formulae for resistance-network analogues solving the finite difference approximation to the partial differential equation

$$\text{div } K \text{ grad } U = g \tag{1}$$

by electrical measurement can be conveniently derived from the transcribed form of equation (1)

$$\sum_{\nu} K_{\nu}\left(\frac{\partial U}{\partial n}\right)_{\nu} \delta A_{\nu} = g_0 \Delta v_0 \tag{2}$$

This will be applied here to the derivation of the design formulae for a resistance-chain analogue for solving equation (1) in cases of spherical symmetry.

Let the space be divided into concentric spherical shells (see Fig. 1); the mean radius of one such shell be r_0, and the distance from the middle of this shell to the middle of the

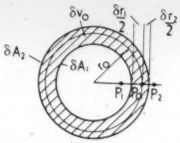


Fig. 1. Volume element Δv_0 (spherical shell of radii $r_0 - \delta r_1/2$ and $r_0 + \delta r_2/2$)

next inner shell, of radius r_1, be δr_1 and to the middle of the next outer shell, of radius r_2, be δr_2. The radii of the shell of mean radius r_0 are then $r_0 - \frac{1}{2}\delta r_1$ and $r_0 + \frac{1}{2}\delta r_2$, and its volume is

$$\Delta v_0 = \frac{4\pi}{3}[(r_0 + \tfrac{1}{2}\delta r_2)^3 - (r_0 - \tfrac{1}{2}\delta r_1)^3] \tag{3}$$

$$\simeq 4\pi r_0^2(\delta r_1 + \delta r_2)[r_0 + \tfrac{1}{2}(\delta r_2 - \delta r_1)] \tag{3a}$$

At the outer surface of the shell,

$$\left(\frac{\partial U}{\partial n}\right)\delta A = 4\pi\left(r^2\frac{\partial U}{\partial r}\right) = -4\pi\frac{\partial U}{\partial(1/r)},$$

which is approximated by

$$-4\pi(U_2 - U_0)\Big/\left(\frac{1}{r_2} - \frac{1}{r_0}\right) = 4\pi(U_2 - U_0)r_0r_2/(r_2 - r_0);$$

similarly, at the inner surface of the shell

$$\left(\frac{\partial U}{\partial n}\right)\delta A \simeq 4\pi(U_1 - U_0)r_0r_1/(r_1 - r_0).$$

* LIEBMANN, G. *Brit. J. Appl. Phys.*, **5**, p. 362 (1954).

Hence, the required difference approximation to equation (1) becomes, for problems of spherical symmetry, with K_1 and K_2 being the values of K at the inner and outer surfaces of the shell (or mean values, averaged over the volumes between r_0 and r_1, and r_2 and r_0)

$$\frac{K_1r_0r_1}{\delta r_1}(U_1 - U_0) + \frac{K_2r_0r_2}{\delta r_2}(U_2 - U_0) = \frac{\Delta v_0}{4\pi}g_0 \tag{4}$$

In more general notation, where the suffix m denotes the mth spherical shell, counting from the centre of the sphere, such that r_m takes the place of r_0, etc., this difference equation reads

$$\frac{K_{m-1,m}r_{m-1}r_m}{r_m - r_{m-1}}(U_{m-1} - U_m) + \frac{K_{m,m+1}r_mr_{m+1}}{r_{m+1} - r_m}(U_{m+1} - U_m) = \frac{\Delta v_m}{4\pi}g_m \tag{5}$$

$$= \frac{1}{4}r_m(r_{m+1}^2 - r_{m-1}^2)g_m \tag{5a}$$

The relation (5) is satisfied by the resistance chain, Fig. 2, where the current I_m is fed in at the junction point P_m of the

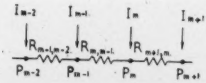


Fig. 2. Resistance-chain analogue for the solution of problems of spherical symmetry

resistances $R_{m+1,m}$ and $R_{m,m-1}$, if the resistance $R_{m,m-1}$ between points P_m and P_{m-1} has the value

$$R_{m,m-1} = \frac{r_m - r_{m-1}}{K_{m,m-1}r_{m-1}r_m}R_0 \tag{6}$$

R_0 being a suitable value of resistance (an apparatus constant), and the current I_m has the value

$$I_m = \frac{\Delta v_m g_m}{4\pi R_0} = \frac{r_m(r_{m+1}^2 - r_{m-1}^2)g_m}{4R_0} \tag{7}$$

Formulae (6) and (7) apply equally whether a uniform increase in r is assumed, or not; in particular, it is possible to use finely graded intervals in some part of the resistance chain, and coarsely graded intervals in other parts. A grading which has been found useful in some problems is one in which the space intervals increase in geometric progression, $r_m = aq^m$, where a represents "unit radius" at $m = 0$, and q is a suitable numerical factor larger than 1. Then $R_{m,m-1} = (q-1)R_0/(aK_{m,m-1}q^m)$, etc. It is particularly easy in this grading to represent an "infinite" medium, for which $g = 0$, and $K =$ constant. Suppose that the grading in geometric progression has been carried out up to a point P_M, at equivalent radius $r_M = aq^M$, where $K = K_M$. Summing all resistance values $R_{M+1,M} + R_{M+2,M+1} + \ldots$, one finds that the resistance of the whole chain from point P_M to "infinity" becomes $R_{0i}(aK_Mq^{M-1})$; it is to the free terminal of this resistance that the conditions prescribed for the function U at infinity are applied.

The author wishes to thank Dr. T. E. Allibone, the Director of the Laboratory, for permission to publish this note.

Research Laboratory, G. LIEBMANN
Associated Electrical Industries Ltd.,
Aldermaston, Berks.

New books

Memorandum on gamma-ray sources for radiography. 2nd Edition. (London: The Institute of Physics, 1954.) Pp. 28. Price 3*s.* 6*d.*

This memorandum is the second edition of one which was first published in 1952. The subject matter has been brought up to date and has been extended in many respects, particularly as regards the number of radioactive materials that receive consideration. The preparation has again been in the hands of a Committee of the Non-destructive Testing Group (formerly the Industrial Radiology Group) of The Institute of Physics and the intention has been to provide a brief statement on gamma-ray sources for radiography, particular reference being made to the products of the Atomic Energy Research Establishment (A.E.R.E.), Harwell, Berks, and the Radiochemical Centre (R.C.C.), Amersham, Bucks. By means of these sources, a wide variety of useful radiography can be carried out without the need for X-ray equipment. This is only fully realizable if the physical characteristics of the radioactive materials are appreciated, and this memorandum summarizes such information as is of importance in industrial radiography. Attention is also given to radiographic technique, protection of personnel and the costs involved. An appendix is added, in which summaries of much of the information are given.

Proceedings of the first International Congress of Electron Microscopy, Memoire hors série no. 1. (Paris: de la Revue d'Optique Théorique et Instrumentale, 1953.) Pp. 768 (paper covers). Price 8000 francs.

This large volume gives the text and illustrative micrographs of some one hundred papers read at the Electron Microscopy Congress in Paris in 1950 and published in 1953. The volume includes the full text of a number of papers from Japanese authors read in title only at the meeting: this is hardly satisfactory as discussion was not possible. A very wide range of subjects is covered, showing clearly the impact of the new microscopy in diverse fields. The broad distribution of papers by subject is as follows: applications in chemistry 19, in metallography 11 and in biology 30. Electron diffraction has 10 papers and electron optics 35.

General comments only are permissible in a review of such a collection of papers: perhaps the points may be made that the high definition electron micrographs and some of the fine electron diffraction patterns deserve fine screen reproduction on art paper: while the majority of the papers are novel and of great interest, a few are on well worn themes and scarcely deserve inclusion: a book which will be referred to for many years should have hard covers: the price puts the book beyond the reach of most individuals: authors' replies to discussion are omitted in some cases.

Summing up, this is an important volume which will find a place in all laboratories concerned with the microstructure of matter and in many more general libraries, but it is a matter for regret that so few individual research workers will be able to afford it.

F. W. CUCKOW

High voltage laboratory technique. By J. D. CRAGGS, M.Sc., Ph.D., F.Inst.P., and PROF. J. M. MEEK, D.Eng., F.Inst.P. (London: Butterworth's Scientific Publications, 1954.) Pp. ix + 404. Price 65*s.*

Twenty-five years ago, the number of laboratories in the world devoted to high voltage research was not more than a dozen, and the classic treatises of Peek on high voltage engineering and Roth on Hochspannungstechnik dealt mainly with the basic engineering phenomena important at high voltages. Since then, many universities have built special laboratories for the teaching and study of these phenomena, and nuclear science has drawn heavily upon many ingenious high voltage sources. A textbook on modern techniques has been needed, and Professor Meek and Dr. Craggs have done a great service by meeting the need with this comprehensive book. Naturally, they have drawn heavily upon the scientific output of the High Voltage Laboratory of the Metropolitan-Vickers Electrical Co. Ltd., where they worked for a decade, but the book covers the ground very thoroughly, as reference to the extensive bibliography at the end of each of the eleven chapters will show.

A sign of the times appears in the order of the chapters. The generation of high voltage direct current, relatively unimportant in 1930, is given pride of place in Chapter 1, mainly because of the various sources of direct current required for ion acceleration. This has resulted in far more work being done on the stabilization and the voltage measurement of high direct voltages, reported in Chapter 2, than of alternating voltages; they can be measured to an accuracy of the order of 0·01‰ or even 0·001‰, whereas accuracy of alternating voltage measurement is not better than one in a thousand.

The generation of impulse voltages and their measurement by sphere gaps, potential dividers and oscillographs is fully covered in five chapters, and the basic mathematical papers on this subject have been well presented in very clear form for the student, thus saving a great deal of back reference to the classic papers on this subject. The chapter on the generation of square high voltage pulses recounts for the first time in book form (apart from an American publication) the war-time effort to produce high power pulsed sources for magnetrons and is a subject of increasing importance. There is also a chapter on the generation and measurement of very high currents up to half a million amperes.

A careful reading of the book fails to reveal any glaring omissions. The book will be a valuable reference for all workers in the high voltage field of study, as well as an excellent textbook for all electrical engineering students.

T. E. ALLIBONE

Nuclear moments. By NORMAN F. RAMSEY. (London: Chapman and Hall Ltd., New York: John Wiley and Sons, Inc., 1954.) Pp. x + 169. Price 40*s.*

This book is based on a section of *Experimental nuclear physics, Vol. I*, edited by E. Segre, and has been issued separately to serve as a textbook on nuclear moment measurements for those chemists and solid state physicists wishing to use this powerful new tool of research. First, brief derivations of the energy splitting introduced by electric quadrupole and by magnetic dipole moments are presented, and ample references to original, more detailed treatments are given. The next chapter is devoted to the various methods of measuring the level splitting, ranging from optical spectroscopy through molecular beam methods to microwave spectroscopy. Up-to-date tables of results for many isotopes, with brief discussion of their significance, constitute another chapter, and the final chapter is headed chemical and solid state applications. This chapter covers an extremely wide variety of measurements and their interpretation, but since each occupies only two or three pages, there is little or no critical discussion of the possible errors of interpretation.

The many references to original work will prove invaluable to workers wishing to take up any of the methods discussed, but the reviewer cannot help feeling that most of the solid state physicists and chemists to whom the book is addressed would have very much appreciated a section stating clearly just what information can be obtained by which methods. It is the lack of such a summary which makes this final chapter appear so disjointed. The list of about 1000 references ensures the book an honourable place as a literature guide.

W. M. LOMER

The physics and applications of secondary electron emission. By DR. H. BRUINING, M.B.E. (London: Pergamon Press Ltd., 1954.) Pp. xii + 178. Price 25*s*.

Whether or not readers have studied the author's book, *Die Sekundärelektronenemission fester Körper*, published in 1942, they will appreciate this new monograph, which is roughly divided into two parts concerned respectively with the physics and with the applications of secondary emission.

A valuable feature of the first section is the comprehensive collection of secondary emission yield curves for metals and metal compounds. The underlying theories which have been advanced to account for secondary emission are also discussed in an admirably clear way, but the quantum theories might perhaps have been described more extensively. The author's elegant "universal" yield curve, although undoubtedly well based, appears to need adjustment at high relative voltages, and would justify further examination.

The section on the varied technical applications of secondary emission is important and stimulating, although in places necessarily somewhat condensed. The process of condensation in the description of output tetrodes has been carried to an excessive degree, while the author's comments on C.P.S. pick-up tubes are somewhat surprising, but in general there is little to mar a work in which errors and blemishes are so commendably absent.

S. RODDA

The sun, the solar system, Vol. 1. Edited by GERARD P. KUIPER. (London: Cambridge University Press; Chicago: The University of Chicago Press, 1954.) Pp. xix + 745. Price 94*s*.

This is the first volume of a large-scale project intended to cover the physics of the solar system. I use the word physics deliberately, because the book is primarily a physicist's reading; it is unlikely to appeal to those astronomers who do not care for discussion in terms of the most recent experimental and theoretical physics.

The volume is a collection of up-to-date monographs by twenty-three leading American, British and Continental workers; the chapters can be selected according to the tastes of a wide range of physicists who might stray into solar topics, and will attract many besides those occupied in each line of research. Although the data and the references emphasize problems of to-day's interest, the writers seem to have been briefed to introduce their subjects historically. Much of the book provides invaluable material not elsewhere available even in the technical journals; for instance Professor Cowling's 60 pages on solar electrodynamics will stand for many years as a critique of what can and what cannot be accomplished with classical electromagnetic theory when applied to the motions of an ionized gas on the large scale. Professor Minnaert's analysis of spectral line contours and the opacity of radiating and convected solar material is another monograph comparable with Cowling's, as exhibiting criticism which has evolved in the mind of a master and is an essential safeguard in future research.

Experimental physicists on the look-out for techniques applicable in other regions will enjoy especially the chapters on photometric and spectrographic devices, ultra-violet rocket cameras, radio-frequency reception aerials, the extraction of sunspot fields by polarization optics on a delicate and very minute scale, and the Babcocks' double photomultiplier circuit for scanning the solar disk and mapping its magnetic field. But experimenters hoping to "pick up something" may be warned that these very readable accounts cover a background of the most consummate skill in trial and error over many years.

Goldberg's introductory essay includes a history of solar research, and brings a sense of proportion which writers in other branches of physics would do well to emulate. The most purely astronomical chapter is Strömgren's assessment of stability and energy sources which classifies the sun among evolving stars. It is significant that perhaps the most detailed treatment of any portions of the sun is van de Hulst's 120 pages on the physics of chromosphere and corona, topics whose mysteries had to be briefly glossed over in earlier solar literature.

There can be few recent works more stimulating to the practising physicist to read, or more useful to the specialist who will take it into regular use.

MARTIN JOHNSON

Wärmespannungen infolge stationärer Temperaturfelder. By E. MELAN and H. PARKUS. (Wien: Springer-Verlag, 1953.) Pp. v + 114. Price 31*s*. 6*d*.

The authors have set out the basic equations for the calculation of stresses due to temperature fields, and shown the methods of solution in a large number of cases of practical interest. Attention is restricted to cases where the body is made from material which obeys the ordinary laws of elasticity. Variable-state problems, and cases where the material shows plasticity in greater or lesser degree, are promised for a future volume. Only elementary mathematics are used, but are made to suffice for many problems of practical interest.

The production is of the high standard we are accustomed to from Springers.

J. H. AWBERY

The statistical approach to X-ray structure analysis. By V. VAND and R. PEPINSKY (Pennsylvania, U.S.A.: The Pennsylvania State University, 1953). Pp. xvi + 98. Price $1.50.

In 1953 the American Crystallographic Association published a monograph entitled *The solution of the phase problem* by Hauptmann and Karle. The problem is the determination of the relative phases of the X-ray beams diffracted by a crystal, and its solution would be a momentous event: direct images of crystal structures could be formed; the present methods of structure determination could be dispensed with and machines could take over all problems.

The present book shows that this claim is invalid. The method proposed has been shown to fail on a simple hypothetical structure, and theoretical reasons are given for this failure. The success that Hauptmann and Karle had with naphthalene is shown to be exceptional, and crystallographers can rest assured that there will be room for their ingenuity and ability for some time to come. The authors have added some further considerations of their own, but they do not exhibit any great confidence that a general solution of the phase problem is yet possible for a complicated structure.

The book is reproduced from typescript and is bound in thin card. It is obviously not intended as a permanent work

but it is certainly a useful contribution to the subject in its present state. H. LIPSON

The structure of metals and alloys. 3rd Ed. By W. HUME-ROTHERY and G. V. RAYNOR. (London: The Institute of Metals, 1954.) Pp. viii + 362. Price 35*s*.

The subject of metals is not a tidy one. Its history is full of chance discoveries, and generalizations have emerged only after much searching in most unlikely places. The authors of this book have been prominent in such work, and it is a tribute both to them and to the book that new editions have been steadily required for nearly twenty years.

The opening section—on the nature of interatomic forces—is extremely good, and should appeal particularly to the non-mathematical reader who wishes to learn something of such concepts as π-orbitals, hybridization and so forth. Sections on the structures of the elements, on primary solid solutions and on intermetallic compounds then follow and contain the "meat" of the book.

The introduction of two final sections—on ferrous alloys and on crystal imperfections—provides the main difference from earlier editions. These sections, however, lack something of the original treatment found in the other sections; the authors show a tendency to "sit on the fence" when a clear lead might be expected. For example, they describe Wood's work as "interesting" but omit to mention that of Warren, Boas, Hirsch and others which has shown Wood's views to be wrong.

One would have wished that the kX unit had been dispensed with; it was intended only as a unit of wavelength measurement and in this book practically every measurement quoted could have been in Ångstrom units, with greater clarity for the student.

The production of the book is good, although occasional figures (e.g. 9 and 10) are poor. Footnotes are used much more than is considered good modern practice.

These details apart, the book can be confidently recommended to the practically inclined as a valuable link with modern theories of the metallic state. H. LIPSON

Elasticity, plasticity and structure of matter. By DR. R. HOUWINK. (London: Cambridge University Press, 1954.) Pp. xviii + 368. Price 45*s*.

This is the second edition of Houwink's well-known book, corrected and, the publishers assure us, brought up to date. It is an unqualified disappointment to the reviewer. With the exception of some half-dozen trifling additions to the footnotes, the first 205 pages are identical with those of the first edition. The remainder of the book is similar except that 17 pages on the molecular structure of rubber have been replaced by 7 pages of new material, and about 9 other pages on artificial rubber and proteins have been revised. No rheological book, however good, is worth reprinting unless it is revised to comprehend recent developments; this reissue virtually ignores all literature later than 1937.

V. G. W. HARRISON

Zirkonium, seine Herstellung Ergenschaften und Anwendungen in der Vakuumtechnik. By Dr. W. ESPE. (Füssen: C. F. Winter'sche Verlagshandlung, 1954.) Pp. 74. Price D.M. 5.40.

The first half of this paper-backed pamphlet contains a brief account of the preparation of various grades of zirconium and of their general properties (thirteen pages), a very short account of the working of the metal with a statement of the precautions necessary when using the metal powder, and an account of the interactions of the metal with gases (nine pages). This is followed by fifteen pages devoted to typical methods of using the metal as a getter, particularly in vacuum tubes.

There is little in Espe's booklet that cannot be found in Miller's recent general monograph on zirconium which is likely to be more readily used in this country. However, at the end of Espe's monograph the analyses of several grades of zirconium, the general properties of the metal, its resistance to attack by various chemicals and the method for purifying commercial zirconium powder are conveniently tabulated. Recipes for ten preparations of zirconium powder as it is applied in gettering are also tabulated at the end of the book and this may prove to be a useful compilation.

G. A. GEACH

The present state of physics. Arranged by F. S. BRACKETT. (Washington: American Association for the Advancement of Science; London: Bailey Brothers and Swinfen Ltd., 1954.) Pp. vi + 265. Price 60*s*.

This book is the published version of a symposium which is, in fact, a collection of seminars. The title is too ambitious for a short collection but will probably not greatly mislead since no one would expect a single short volume to contain other than a brief selection of topics. Indeed, the purpose of the title is presumably to indicate that the selection has been made from the whole range of subjects in which physicists are actively interested. The broad sections of physics represented in this volume are certainly those which are of first interest to-day, namely, elementary particles, the solid state, physical chemistry and biophysics. Furthermore the contributors are leading experts whose opinions command respect.

On the other hand the articles are unequal in that some essay a summary of a whole subject whereas others give condensed accounts of recent original researches of the author and his colleagues. The book opens with a short note on the magnetic moment of the electron by P. Kusch; it is followed by E. P. Ney's account of the work of his group on cosmic ray experiments at high altitude with some emphasis on heavy particles. J. C. Street surveys more generally progress in the study of cosmic radiation, and is concerned with processes to which the incident radiation gives rise, including, naturally, meson theory.

In the section on solid state Lark-Horovitz provides a pretty comprehensive review of the theoretical and experimental principles of semi-conductors, with a valuable bibliography of 350 references. Conversely, Bardeen confines himself to a very readable account of transistors. Similarly von Hippel deals with the researches on the ferroelectric properties of barium titanate carried out under his direction.

The section on physical chemistry comprises an article by P. J. W. Debye mainly on the optical scattering method of investigating long chain polymers and a review, with a bibliography, of chemical kinetics in some biological systems by R. Lumry and Eyring. This leads naturally to biophysics which is represented by a partly original paper by F. Brink on the mechanism of communication in nerve cells and a review by F. H. Johnson on the use of luminescence in finding the way in which reaction rate is controlled in living organisms.

There is no doubt that all these articles are interesting and valuable to anyone wishing to keep abreast of the topics treated, but they have neither a consistent objective nor a consistent level of appeal. For this reason the price will tend to discourage sales to individuals, although libraries would be well advised to secure material from such expert contributors. S. WHITEHEAD

Notes and comments

Elections to The Institute of Physics

The following elections have been made by the Board of The Institute of Physics:

Fellows: W. E. Duncanson, H. F. Freunlich, W. A. Jennings, D. J. Littler, J. L. Putman.

Associates: K. R. Allen, D. F. Arthur, W. W. Foster, R. V. Hesketh, G. A. Ingram, P. G. E. F. Jones, W. G. Kimmins, A. E. Kiss, R. M. Lee, T. P. MacRae, A. M. Masoud, R. H. Morris, W. E. J. Neal, E. V. Reeves, N. Richards, P. C. Roberts, G. H. Tattersall, G. C. Sims, K. H. Wilkinson, P. H. Wolfenden.

Forty-three Graduates, ten Students and seven Subscribers were also elected.

A further volume of laboratory and workshop notes

A third selection of laboratory and workshop notes and cognate articles reprinted from the *Journal of Scientific Instruments* has just been published, with the title *Laboratory and Workshop Notes 1950–1952* by Edward Arnold (Publishers) Ltd., at 30*s.* net. Once again through the courtesy of the authors, the royalties from the sale of this volume will be placed to the credit of The Institute of Physics' Benevolent Fund.

The new collection is divided into the following sections: Mechanical design; Laboratory tools, methods and hints; Optical and thermal devices and techniques and infra-red spectrometry; Devices for liquids and gases; Vacuum devices and techniques; Electrical devices and techniques.

One reviewer referred to the previous volume as "a catalogue of ingenuity" and this might equally be said of this new selection. Indeed, it was because of the success of the first two volumes that the Board of The Institute of Physics decided to publish further volumes from time to time in order to assist laboratory workers, and at the same time to help its Benevolent Fund.

For these excellent reasons, we invite our readers at home and oversea to order a copy of this new volume, if indeed they have not already done so.

The jubilee of the invention of the thermionic valve

In connexion with the celebration of the jubilee of the invention of the thermionic valve by Sir Ambrose Fleming on 16 November, 1904, a conversazione will be held in the Electrical Engineering Department of University College, London. A plaque commemorating the occasion will be unveiled by the Lord President of the Council. In addition to exhibits and documents relating to Sir Ambrose Fleming's work while Professor of Electrical Engineering at the College, examples of recent researches will be on view.

Admission to the conversazione, which will extend over the three days 16–18 November, will be by invitation ticket only. A limited number of tickets for the 17 November, when the exhibition will be open from 3 to 10 p.m., is available on application to the Assistant Secretary, University College, Gower Street, London, W.C.1.

The British Standards Institution

The growth of the B.S.I. in recent years is well illustrated in some figures given in its *Annual Report* for the year ended 31 March, 1954, which has just been published as a 200-page book. No less than 286 new standards were issued and the 2388 Standing Committees on which 14 000 persons served held 3470 meetings during the year; and just under one million copies of British Standards were sold in the same period. There are some sixty "Industry" Committees controlling the work of the numerous technical and sub-committees, including one for the scientific instrument industry; summaries of their work are given. An encouraging part of the report is the record of an increasing amount of international agreement, especially on symbols, terminology, values of constants and so forth.

The *British Standards Yearbook* (price 12*s.* 6*d.*) lists the 2500 current British Standards and other relevant data and has a brief description of the subject-matter of each of them. A comprehensive index makes this a useful reference book.

Copies of both of these publications may be obtained from British Standards House, 2 Park Street, London, W.1.

Journal of Scientific Instruments

Contents of the November issue

ORIGINAL CONTRIBUTIONS

Papers

The application of the electrometer valve to charge measurement. By M. J. Morant.
The effect of back-scattering of electrons on measurements in β-spectroscopy and absolute counting. By J. G. Balfour.
An electron diffraction specimen chamber with a specimen vacuum lock. By R. B. Kehoe, R. C. Newman and D. W. Pashley.
A stable and sensitive d.c. amplifier with high input resistance. By S. O. Nielsen and Th. Rosenberg.
A simple vane anemometer giving continuous and direct indication of wind velocity. By A. E. Hawkins.
An autocorrelogram computer. By G. Revesz.
Vibrating plate method of producing powder ridges in a sound field. By Mary D. Waller.
A thermionic ionization gauge of high sensitivity employing a magnetic field. By G. K. T. Conn and H. N. Daglish.
A modified apparatus for measuring the swelling of polymer films in solvents. By N. Stolow.
Temperature control of a large water-bath using a resistance thermometer. By W. P. Hutchinson, E. W. Pulsford and A. G. White.
The characteristics of a quartz-fibre electroscope. By F. Harlen.

Laboratory and workshop notes

A portable integrating circuit for low radioactivity counts. By A. H. Ward.
A packless diaphragm-type valve for use with standard glass chemical pipe-line. By J. A. G. Smith.
A simple pulse-counter with many laboratory uses. By J. R. Harris and C. G. Gurr.
A demountable vacuum seal. By C. B. Richards and J. R. W. Smith.

NOTES AND NEWS

Correspondence

The use of quenching circuits with Geiger–Müller counters. From E. H. Cooke-Yarborough: R. L. Gordon.
A portable high-speed cathode-ray oscillograph. From L. S. Allard.
The influence of the ballast resistance on the performance of Penning vacuum gauges. From G. K. T. Conn and H. N. Daglish.
Design of double-beam spectrometers using multiple monochromators. From J. Pliva.

New instruments, materials and tools
Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique).
All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions.* (Price 2*s.* 6*d.* including postage.)
REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.
SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.

Summarized proceedings of a conference on X-ray analysis—London, November 1953

The Autumn Conference of the X-ray Analysis Group of The Institute of Physics was held in London on 20–21 November, 1953. The first day's session was devoted to the subject of film and counter methods and the second day's to identification by X-ray analysis. The papers presented and the discussion on them are summarized in this report.

The first paper was given by DR. R. H. HERZ (Kodak Ltd., Harrow) and was entitled "Survey of the photographic fundamentals of X-ray films." He felt that a thorough understanding of the inherent properties of X-ray films could enable a better use to be made of their potentialities. The response of emulsions to X-rays could best be understood by reference to the mechanism of latent image formation and the widely-accepted Gurney–Mott conception of this process[1] was outlined. The basic mechanism for light and for X-rays is regarded as similar, but differences are found between the number of light and X-ray quanta required to render a single silver bromide grain developable; several light quantum hits are required, whereas only one X-ray quantum needs to be absorbed. For the initial part of the density *versus* exposure curve (*D*–*E* curve), the density (i.e. the number of developable grains) is directly proportional to the number of X-ray quanta absorbed per unit area. When this latter number exceeds a certain value, so that the probability of several grains being hit twice or more is increased, the *D*–*E* curve flattens out. Some of the fine-grained X-ray films show a linear *D*–*E* relationship up to $D = 1{\cdot}8$.

The speed of an X-ray film may be defined as the value of D/E at a given D above fog level. Thus, of the three films whose *D*–*E* curves are shown in Fig. 1(*a*), film 1 is the fastest. Whilst their *D*–log *E* curves [see Fig. 1(*b*)] are shifted relative

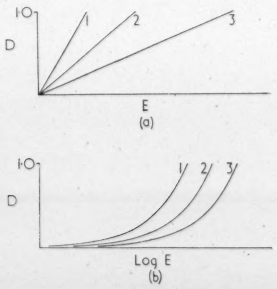


Fig. 1. Density-exposure and characteristic curves of films having different speeds

to the abscissa, their shapes are identical; this means that the inherent contrast [of which the $dD/d(\log E)$ value is a measure] of the three films is the same. With regard to the behaviour of the films with change of wavelength it is found that the speed varies considerably, the efficiency of response being relatively high for the softer radiations.[2] This higher efficiency is not fully realized in practice, however, because this radiation is more readily absorbed in the front emulsion layer. The inherent contrast of a given film is the same for all wavelengths less than 2 Å and is lower for greater wavelengths, for the same reason. This independence of contrast with wavelength is of great advantage when rotating-sector, calibration patterns are recorded, as broad, uniform, unfiltered beams of all but the softest radiations can be used.

The developing conditions have a marked influence on the speed and inherent contrast of the film. Fig. 2 represents a family of characteristic curves of an X-ray film developed for

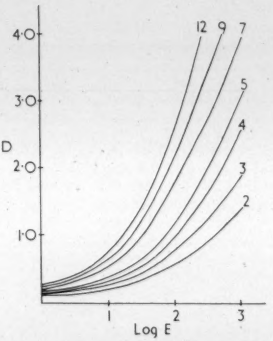


Fig. 2. Characteristic curves as function of developing time in minutes at constant temperature

different times at a constant temperature, showing that the speed, inherent contrast and fog level all increase with increasing development time. The optimum development time for a given film is derived from the maximum contrast and highest speed combined with a tolerable fog value.

A simplified explanation was given for the validity of the reciprocity law for X-ray film.

DR. P. B. HIRSCH (Cavendish Laboratory, Cambridge), in the next paper, "The requirements of X-ray films in diffraction work," considered it useful in discussing the characteristic properties of films to divide the experimental investigations for which films are used into three groups.

The first group is that in which the positions of peaks are required, as in ordinary powder work. Here, shrinkage is the most important factor, particularly when the highest accuracy is needed in the determination of the unit cell dimensions. Whilst it can be minimized, it cannot easily be avoided, and all accurate determinations of interplanar

spacings or cell dimensions are carried out by methods calculated to eliminate errors arising from uniform shrinkage. The shortest possible time for effective washing is recommended. American workers have found that when film strips, 2 × 12 in., were dried by hanging vertically, the lower end suffered greater shrinkage. They found that this non-uniform shrinkage could be lessened by suspending the strips in a draught-free atmosphere with their lengths horizontal and their plane vertical.[3]

What constancy in the shrinkage effects is required for the accurate determination of cell dimensions? For the van Arkel type of film mounting, it can be shown that

$$\frac{da}{a} = -\cot\theta\left[\phi d\left(\frac{S}{S_K}\right) + \left(\frac{S}{S_K}\right)d\phi\right]$$

where 4ϕ is the angle between the knife-edges, and S and S_K are the distances on the film between the diffraction lines and the knife-edges, respectively. The shrinkage coefficient, e, can be defined as $1 - (S/{}_0S)$, where ${}_0S$ is the true, unshrunk distance between diffraction lines in the camera and S is the distance between the lines on the processed film. To obtain an accuracy of da/a of 1 part in 10^5, it is found that the increment in the shrinkage coefficient due to non-uniformity, f, must be less than 5×10^{-4}, where $f = S/{}_0S - S_K/{}_0S_K$ (${}_0S_K$ and S_K have the same meanings as before, except that they represent the distances between knife-edges). It appears from the experiments mentioned above that f can have values larger than this, so that the accuracy of determination of cell dimensions is reduced. The main trouble about this quantity is its uncertainty, which can only be removed by producing a series of marks of known separations on the film. It would, nevertheless, be helpful if manufacturers could recommend films particularly suitable for accurate cell dimension work, and provide some idea of the magnitudes of e and f likely to be encountered under standard processing conditions.

The second group of experimental investigations concerns those in which weak diffraction lines are to be detected. If the absolute intensity of a line is low, but its intensity relative to background is high, then the film speed is the important factor. For double-coated Ilfex G, about 53% of incident Cu Kα radiation is absorbed by the emulsion layers and 12% by the film base so that there is little likelihood of securing a further substantial increase in speed for soft radiations. For harder radiations, like Mo Kα, approximately 20% is absorbed by the emulsions, so that there are good prospects here for faster films. Whilst a five-fold gain in speed can be had by using zinc sulphide screens, accurate intensity measurements cannot then be made.

When the intensity of a line relative to background is small, (see Fig. 3), then the film contrast becomes important. It

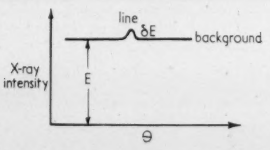


Fig. 3. Detection of a weak line on a strong background

can be shown theoretically that if the film is used in the linear region of the D–E graph, and if the accuracy of light measurement in the photometer is of the order of 0·5%, then the smallest $\delta E/E$ which can be measured is approximately $0{\cdot}01/D$. It follows that the greater is D, the smaller is the relative intensity of a line which is measurable; i.e. the higher the contrast. A microdensitometer record was shown of part of a Debye–Scherrer photograph of an anthracite taken with crystal-reflected Mo Kα on Ilfex Industrial G. By using high D values, but still keeping well in the linear region, it was possible to detect bands whose intensities were only 1–2% of the background. The most important factor influencing the detection of weak bands on a strong background is found to be density fluctuations due to impurities in the film and to development stains. In addition to emulsion artefacts, fast films tend to show very high fog levels after storing for only a few months.

Measurements of intensity by either visual or photometric methods represent the third group of investigations. For single crystal work a constant sensitivity over large film areas is particularly important. Since some form of film calibration is usually employed, a wide range of linear response is not essential; the fastest films are the most suitable, provided the fog level and the density of impurity specks are low. The graininess of films only becomes important when very fine lines are involved, i.e. when the photometer slit must be reduced to about 10 μ. When naked films cannot be used, the rippled background caused by the uneven texture of the black paper becomes objectionable. A better light-shielding material is greatly needed.

Finally, there is also a real need for a fast, single-coated film for certain tasks, such as determining line positions with focusing cameras and counting spots on spotty rings.

Mr. H. S. Peiser (Hadfields Ltd., Sheffield) suggested that, as in Dr. Hirsch's own experience, most of the non-uniform shrinkage resulted from uneven drying, failure to notice this effect in Britain could be due to our damper climate which favoured slow, even drying. Thin metal foils were a satisfactory light-shielding film covering, but he preferred to use light-tight cameras. In answer to a further enquiry of Mr. Peiser's, Dr. Hirsch stated that in the American work on non-uniform shrinkage, the films had not been treated with a wetting agent.

Prof. E. G. Cox (University of Leeds) said that he found that two sheets of black paper averaged out texture marks.

Dr. Herz believed that certain infra-red gelatine filters were of very uniform texture and he found that they give adequate protection against dim light.

In reply to a question regarding the possible virtues of chemical intensification processes, Dr. Hirsch said that these were often of value, but that the film lost D–E linearity. However, Dr. Herz stated that Corney and Splettstosser[4] had shown that the linearity can be well maintained.

Dr. R. E. Franklin (Birkbeck College, London) mentioned that in France, the Kodak Company produced single-coated X-ray films as standard practice.

In his contribution, "Microdensitometer design," Dr. C. J. Brown (I.C.I. Ltd., Manchester) said that judging by the amount of literature that had been published on microdensitometry[5] compared with what he had observed in visiting other laboratories, he had come to the conclusion that it must be the most discussed, but least used, technique associated with X-ray crystallography. In the past, the available microdensitometers have been generally unsatisfactory for the purpose they have been supposed to fulfil, good instruments have been beyond the reach of most laboratories and the eye is such a good judge of light intensity that frequently the use of an instrument has not been regarded as worth while.

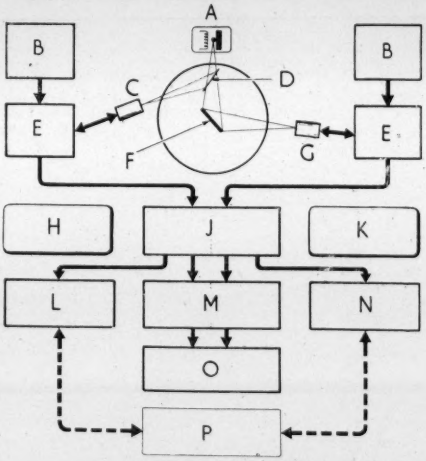
There are between 20 and 30 different published designs of microdensitometers. The majority of these are intended mainly for the densitometry of Debye–Scherrer photographs and are direct deflexion instruments. The chief defects are those arising from the difficulty of keeping the light source and photocell response constant, of providing convenient and satisfactory means for accommodating and manipulating long strips of film, and in overcoming slowness or awkwardness in operation. All of these difficulties have been overcome in the manually-operated instrument described by Taylor.[6] This is a development of the original microdensitometer of Dobson[7] who devised the split-beam principle using a single light source. One beam passes through the X-ray film and the other through a neutral grey wedge, and, by means of a shutter lever, the two beams are allowed to fall alternatively on one photocell. The neutral wedge is moved until the galvanometer deflexion is the same for both circuits, and at the balance point the wedge reading is directly proportional to the film density, and therefore to the X-ray intensity over the linear portion of the *D*–*E* curve. The provision of a clicking ratchet device on the drum of the film traversing screw to mark equal intervals of, say, 0·01 mm, greatly increases the ease and speed of operation. The controls for moving the film, wedge, and shutter are on the left side of the instrument, leaving the right hand free for recording the measurements.

In general, an instrument of this design is not entirely suitable for measuring single-crystal reflexions where the integrated intensity over the whole spot is required. Various suggestions for performing this have been published, amongst them the unusual method of Astbury, requiring a source of α-rays from polonium[8] and the integrating device of Robinson[9] which is impracticably tedious for modern organic crystal structure determinations. The most successful instrument so far appears to be the scan photometer of Robertson and Dawton.[10] A scanning instrument, developed along the same lines, was on show in the exhibition of apparatus. In this system the diffraction spot is projected on to a rectangular aperture on a screen, and a scanning drum rotating behind this aperture enables the light transmitted through each element of the spot in turn to fall on a photomultiplier tube. The current from this is amplified and fed to a galvanometer having a period much slower than the scanning time, so that the integrated intensity is recorded. The null-balance optical system is not used, but other devices are used to ensure reasonable constancy in operation. An anti-logarithm circuit is included to extend the linear range and means are provided to ensure that zero scale reading is obtained when there is no X-ray spot in view. A triple objective nosepiece is used to allow different spot sizes to be projected into more or less the same area, so that the instrument sensitivity is not altered correspondingly. One difficulty not yet overcome arises from the variable background of Weissenberg photographs.

Dr. Brown felt that the whole subject of intensity measurement of single-crystal reflexions now needed thorough investigation so that comparisons could be made between visual assessment and photocell densitometry, or between the Robertson–Dawton, Wiebenga,[11] and counter diffractometer methods.[12]

Following a question on the meaning of optical density by Mr. E. J. W. Whittaker (Ferodo Ltd., Chapel-en-le-Frith), Sir Lawrence Bragg said that it might be useful to re-examine the method proposed by Brentano,[13,14] in which the light scattered by the silver grains of the emulsion is measured. The photocell current is a direct measure of exposure and not a logarithmic function. The method is more accurate at low densities than the conventional transmission method and is not sensitive to emulsion discolorations, etc.

The relative merits of film and counter diffractometer methods were next discussed by Dr. R. L. Gordon (Safety in Mines Research Establishment, Sheffield). In his laboratory both techniques are in use in routine quantitative and semi-quantitative analyses. In such applications the much higher capital cost and maintenance difficulties make counter methods worth while only where the volume of work is considerable. For straightforward identification of unknown materials Dr. Gordon finds photographs easier to recognize at a glance than chart records obtained with a diffractometer, although a number of speakers in the subsequent discussion disagreed with this assessment. For the differentiation between samples whose probable constituents are known and where the positions and intensities of only a few lines have to be measured with accuracy the diffractometer is both quicker and more convenient. It is best to measure the intensities of a Debye–Scherrer line using a scaler, the background being established from a chart record. Since the absorption of a sample can readily be measured on a diffractometer there is often no need for the use of internal standards mixed with the specimen.[15] Dr. Gordon's instrument is shown schematically in Fig. 4. The direct



(Crown copyright reserved)

Fig. 4. Schematic diagram of Safety in Mines Research Establishment counter diffractometer

A, X-ray tube; *B*, counter power supply; *C*, monitor counter; *D*, Al foil; *E*, quenching unit; *F*, specimen; *G*, main counter; *H*, monitor channel; *J*, input and test unit; *K*, main channel; *L*, monitor scaler; *M*, two-channel ratemeter; *N*, main scaler; *O*, self-balancing pen recorder; *P*, timing and control unit.

beam is monitored by a second counter.[16–19] Either scalers or rate meters can be used in the two channels; in the latter case a self-balancing pen recorder is used as a

ratiometer.[20] Quenching circuits[21] are employed both to improve stability of operation of the counters and to make the accurate computation of counting losses possible.

MRS. I. H. HARDWICH (Metropolitan Vickers Electrical Co., Manchester) found that diffractometer charts were no more awkward to assess than films. In her experience stabilization of the X-ray tube output was preferable to monitoring. She operated her counters without a quenching circuit but with a considerable degree of amplification; consequently she could ignore counting losses for rates below 1000 counts/s.[22] Dr. Gordon pointed out that while counting loss corrections could sometimes be neglected, especially when the counter is operated in the proportional or semi-proportional region, there is no truly linear region since some pulses are invariably lost at all counting rates.

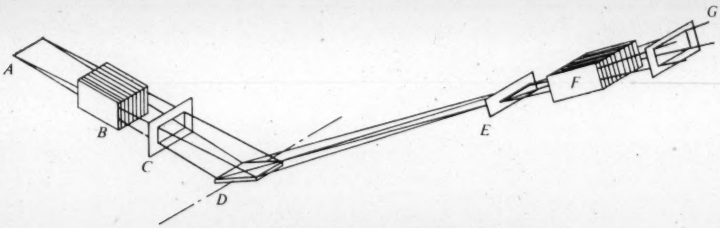
DR. W. A. WOOSTER (University of Cambridge) observed that for measurements of very weak intensities counter methods were preferable to photographic techniques.

DR. G. D. RIECK (Philips Ltd., Eindhoven) described the design and discussed some of the applications of the Norelco counter diffractometer. This precision industrial instrument[23] was designed by Dr. W. Parrish of the North American Philips Laboratories and some 300 diffractometers are now in use. The instrument is used with a standard Philips X-ray tube whose anode end is shaped to allow the counter to reach a back-reflexion position of 165° in 2θ. The Bragg–Brentano focusing principle is employed; Soller slits allow a long view of the line focus to be utilized so that a high available X-ray intensity is coupled with maximum resolution (Fig. 5). The distance from focal spot to sample is 17 cm and the resolution is such that the Kα-doublet separation of cobalt can be seen at $2\theta = 40°$. The scanning counter which is connected to the sample holder by 2 : 1 gearing can be moved continuously at speeds varying from $\frac{1}{8}$° to 2° per min, or in steps. The angular accuracy is very high, the backlash being less than 0·0025°; the positions of sharp lines can be read off a recorder chart to about 0·01° in 2θ so that great precision is possible in the measurement of lattice constants.

The X-ray output of the generator is kept constant to better than 1% by controlling both high tension and tube current. The electronic circuits provide for direct pulse counting using either a "fixed time" or a "fixed counts" method. Alternatively, for somewhat lower accuracy a counting rate meter and chart recorder can be used, a variety of integrating time constants and recorder speeds being available. An additional accessory has been designed for the automatic recording with constant statistical errors of backgrounds and line profiles. This consists of a robot which sets the counter, switches on the counting mechanism, measures the time required for a preselected number of counts and records it, then moves the Geiger counter 0·01° and repeats the procedure. This facility has also been found of great use for conveniently detecting very weak peaks since a time-consuming scan can be carried out automatically.

Dr. Rieck pointed out that a versatile instrument of this type can carry out most of the functions of the powder camera with greater convenience and accuracy; somewhat more care is needed, however, in specimen preparation. A coarse-grained sample can lead to considerable errors unless it is rotated in its own plane.

DR. U. W. ARNDT (Royal Institution, London) discussed some recent developments in X-ray counters. A high degree of detection efficiency and stability can be achieved with commercial X-ray Geiger counters, especially when used with quenching circuits.[21] However, the long inherent dead times of these counters limit their usefulness, counting rates in excess of 1000 counts/s not being practicable. With proportional ionization counters, and with scintillation counters, on the other hand, the resolving times are so short that counting losses become completely negligible in any normal X-ray application. Proportional counters lend themselves well to an electronic method of monochromatization.[24–26] Since the amplitude of the output pulse is proportional to the energy of the incident quantum a single-channel pulse-size selector can be used to separate pulses due to the characteristic radiation from "white" background. When this method of operation is combined with the use of a β-filter the amount of non-Kα radiation detected is only about 1%.[25] Dr. Arndt had experimented with a number of different designs of proportional counters. An all-glass type with a carbon cathode[27] had been abandoned because it was not well suited for "home" construction and because of the difficulties of exact location in its housing. Sealed-off, metal-body counters, while easy to construct, tend to show long period changes in characteristics and he now uses a continuous-flow counter through which a gas mixture of 98% of argon (from a cylinder) and 2% of carbon dioxide is passed at a steady rate of about 1 c.c. per second.[28] As an alternative for fast counting, scintillation counters[29] would be very suitable since their resolving time can be even shorter than that of proportional ionization counters. The detection



(By permission of North American Philips Laboratories.)

Fig. 5. Norelco diffractometer slit system

A, line focus; *B*, divergence parallel slit assembly; *C*, divergence slit; *D*, specimen; *E*, receiving slit; *F*, receiving parallel slit assembly; *G*, scatter slit.

efficiency even for Mo Kα radiation can easily be made to approach 100%. Difficulties which can be anticipated with this type of counter are the relatively high background counting rate and possible lack of stability. Owing to the small number of photoelectrons liberated at the photocathode of the electron multiplier per incident X-ray quantum—about fifteen for the best phosphors and electron multipliers at present available—wavelength discrimination is not possible for characteristic radiations from metals near copper.

Proportional counters and, to a lesser extent, scintillation counters, require more elaborate electronic circuits than do Geiger counters. However, the standard circuit elements required for all counting applications, scalers and high tension units are being greatly simplified by the introduction of cold-cathode, decade scaling valves[30,31] and corona stabilizer tubes[32] and their increased use may make the greater complexity of additional circuits more feasible. Examples of units of this type[33] designed at the Atomic Energy Research Establishment, Harwell, were shown in the exhibition which was arranged at the conference.

Dr. G. E. Bacon (Atomic Energy Research Establishment, Harwell) reviewed the techniques in use at Harwell for intensity measurements in neutron crystallography.[34] A monochromatic beam of neutrons of a wavelength of approximately 1 Å is obtained by reflecting the collimated beam of thermal neutrons, emerging from a pile, from a single crystal of lead. Since, under experimental conditions, there is always a high γ-ray flux, a proportional counter filled with $^{10}BF_3$ is the most useful detector since this combines a high efficiency with the ability to discriminate against this background. Heavy shielding is necessary to reduce the additional neutron background (Fig. 6) which even so remains at about

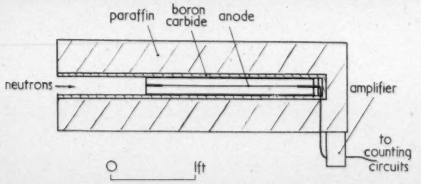


Fig. 6. $^{10}BF_3$ neutron counter with shielding

20 counts per min. The counter is thus very bulky and neutron diffractometers are thus much larger than corresponding X-ray instruments. A pile is a very feeble source of neutrons and the peak counting rates obtained in neutron crystallography are only some hundred counts a minute; very slow scanning speeds are thus necessary and all recording is automatic. Because of the weak intensities available very large samples must be used; in single-crystal work[35] specimens measuring several millimetres are used and extinction corrections are important.[36,37] Absorption effects are generally less serious than in X-ray work. The standard cylindrical powder samples contain about 7 c.c. of material. The resolution of a neutron powder diffractometer is only about a third to a half that of a 19 cm X-ray camera. This is due partly to the lower degree of collimation possible and partly to the wavelength spread of the radiation used; a neutron spectrum contains no sharp lines and the monochromator merely selects a band of wavelengths from the continuous spectrum. Corrections for subharmonics are important only when weak superlattice lines are being investigated. The same basic instrument[38] is used for both powder and single-crystal work; however, in the case of single crystals there is much less chance of overlapping spectra and a much smaller instrument with a lower resolution could be employed if smaller detectors were available. Absolute intensity measurements are easier to make than with X-rays; the neutron scattering factor is independent of the scattering angle.

Dr. W. Cochran (Cavendish Laboratory, Cambridge) pointed out that single-crystal X-ray counter diffractometers could be much smaller than powder instruments, because, just as in the case of neutron diffractometers, a lower resolution could be tolerated. In this type of work[12] the intensities of reflexions can be measured to about 1%, an improvement of about 5 times on the accuracy obtainable by photographic methods. Counter methods are, however, slower and less convenient. Integrated intensity measurements can be carried out more simply by using a stationary crystal in a non-parallel beam from a broad, uniform X-ray tube focus. Several speakers agreed on the difficulty of obtaining this uniformity.

The theme of the Saturday morning session was "Identification by X-ray analysis."

The introductory speaker, Mr. H. S. Peiser (Hadfields Ltd., Sheffield), observed that one of the most impressive features of the X-ray Analysis Group conferences was the balance struck between academic achievements and those applications of X-ray crystallography which have made it an indispensable tool in almost every industrial, government or technological research establishment. Yet, identification, perhaps the most important of the applied crystallographic techniques, had not been fully discussed at previous gatherings of the Group. It was therefore appropriate that an entire session was to be devoted to a discussion of the many experimental methods which have been developed and used in various laboratories. It might well be that just one of these special tricks would make the difference between success or failure in solving someone else's next problem.

Mr. Peiser felt that the crystallographer's first task in industry and technology was to know what was going on in his organization. He must study the raw materials and manufacturing problems. He must know of the quality control requirements and keep an eye open for production processes where uncontrolled variations in conditions may creep in, and where a build-up of unwanted materials may arise. He should bring to bear a deeper understanding of the atomic mechanisms underlying the reactions occurring during production, wherefrom improvements in processes and products are likely to arise. Finally, the crystallographer should seek problems from among the wider research projects that are being investigated in his laboratory.

As to the treatment of samples, about a dozen approaches can be listed. It is often easily possible to separate constituents of a mixture by hand under a binocular, low-power microscope. Some phases may be capable of being selectively dissolved out by water, acids or alkalies. Certain matrices or minor phases may be isolated by electrolytic dissolution, by selective evaporation or by melting. Some mixtures are successfully separated magnetically, others by flotation, by frothing agents or by heavy-liquid, sink-and-float methods, with or without the aid of a centrifuge. When one is faced with a large series of related samples, it is inadvisable to X-ray all the members before looking at one or two of the photographs. Equally, it is a mistake to struggle with the interpretation of one pattern. Frequently,

the comparison of several photographs reveals systematic intensity changes from which the lines can be grouped into those belonging to different phases. Amorphous materials can often be induced to crystallize by careful heating or stretching; occasionally it is possible to infer the presence of an amorphous substance in a mixture by intensity measurements on reflexions from the crystalline constituent. Mixtures of hydrated phases are dealt with best by thermal analysis in conjunction with X-ray photography of the reaction products at various stages. Problems involving solid solution effects are often very awkward, but it is sometimes possible, by judicious chemical change, to convert the unknown material into compounds which do not display an equally extended solid solution series, for example, the oxidation of solid solutions of refractory carbides. A thorough knowledge of the history of the sample makes identification very much easier. Even an incomplete chemical analysis may prove invaluable; a parallel spectrographic test is almost always worth while. Lastly, it should not be forgotten that coarsely grained phases can often be identified more rapidly from their optical characteristics by means of the polarizing microscope.

Whilst the A.S.T.M. index is the most potent method of identification, the most satisfying method is undoubtedly the direct comparison of the test film with a standard material photographed under the same conditions. In every laboratory there is a bias towards certain groups of compounds and the use of only one or two radiations, so it is by no means impracticable to keep a library of standard films. If all else fails, one can still take recourse to Bunn charts[39] or to Ito's method[40] in an attempt to determine cell dimensions for comparison with data available in the literature.[41] This is because many compounds, not yet in the A.S.T.M. index, have been studied by single-crystal X-ray methods.

In presenting his contribution, "Identification of poor reflectors," DR. D. M. C. MACEWAN (Rothamsted Experimental Station, Harpenden) explained that the term "poor reflectors" was to be taken to mean substances which give X-ray patterns which are poor in lines and possibly diffuse. That variety of poor reflectors to which he would mainly confine his remarks, namely clay minerals, represented the only one in which the problem of mere identification had assumed an outstanding practical and theoretical importance. It would perhaps be of interest to those concerned with the general problems of X-ray analysis of mixtures to see how the special difficulties have been tackled by clay mineral specialists.

By definition, the clay fraction of rocks and soils includes all the solid matter of grain size less than 2 μ (effective spherical diameter). This may include material so fine that its X-ray pattern will consist of broad bands, perhaps not even that. If well-crystallized particles of size approaching 2 μ occur as well, it is a very difficult task to *detect* the semi-amorphous or amorphous material by X-rays, let alone identify it, even if it is a major constituent of the mixture. There are admittedly certain means of separating the clay into size fractions, but they do not always allow the amorphous material to be isolated.

The structures of all clay minerals are built up of layers. The linking between layers is much less strong than that within layers, so that these minerals have an easy cleavage, and, in the extreme case, may be readily dissociable into individual free layers. Three types of layer occur; the mica type, as in the hydrous micas, montmorillonite and vermiculite, the kaolin type, as in kaolinite, meta-halloysite and chrysotile, and the chlorite type. The progression from well-crystallized to irregular structure is well illustrated by the groups of minerals which include the completely ordered mineral kaolinite, the randomly ordered mineral meta-halloysite and an intermediate, as yet unnamed, mineral found in fireclays. The X-ray criteria for distinguishing these different members of the kaolin series are shown in Fig. 7. It is seen that although they are distinguishable by the

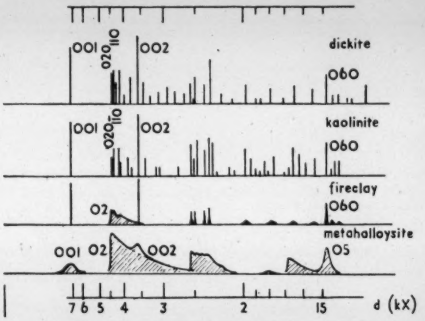


Fig. 7. Powder diagrams of kaolin minerals (after Brindley)

positions of their lines, another criterion, that of line, or band, shape must be introduced. When a pattern of the disordered mineral is measured, the "line position" is an indefinite concept. This is illustrated in Fig. 8. If the measurements are made directly on the photograph, the

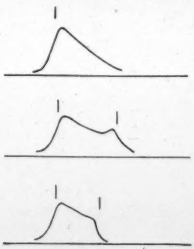


Fig. 8. Visual and photometric assessment of "line" positions. The eye would see "peaks" where the short vertical lines are drawn

location of the "line" is generally judged to be near the position of maximum slope of the intensity curve, whereas with a microdensitometer record or diffractometer curve, the location adopted will generally be that of maximum intensity. Moreover, the shape of the curve and the position of the maximum will change with the crystallite size of the mineral. In some cases, two "lines" are judged to be present by visual examination when in fact there are merely two maxima in a diffuse band, or even a flat plateau with two edges.

A similar transition is observed in the mica-type series, with a completely ordered muscovite or biotite at one extreme and the "turbostratic" montmorillonoids at the other. Again, minerals of this sort are highly reactive and allow various substances to penetrate between the layers and so alter the interlayer separation, giving X-ray patterns of very different appearance. Finally, those clay minerals which are composed of essentially similar layers may form mixtures on any scale,

from a gross mixture of crystallites to an interleaving of different layers with a regular or irregular sequence. The X-ray pattern given by a random interstratification of layers commonly shows a sequence of basal reflexions which, although they may be reasonably sharp, will not form an integral series of orders; they may also be quite diffuse. The interpretation of such patterns may be done in two ways. Starting from an assumed mixture of layers, the type of diffraction pattern may be calculated and compared with that actually obtained, or the "probability" distribution of the layers may be determined directly from the measured diffraction pattern, by the use of a Fourier transform. The latter method, analogous to the determination of the radial distribution of scattering matter in a liquid, gives a direct analysis of the mixture in terms of the interlayer distances, their proportions and the manner of their distribution.

It is clear that for cases like these, the A.S.T.M. card index, as at present arranged, is of very little use. Some of its faults may be corrected, but others are inadequacies in the arrangement itself. Insufficient account is taken of the continuous variability of the position, shape and intensity of the first- and higher-order basal reflexions of certain minerals. Attention is not given to subsidiary identification checks depending on the X-ray behaviour of the material after specific chemical or thermal treatments, although such tests may be essential. Orientation and textural effects, which cannot easily be dealt with, are also of great diagnostic value, allowing a separation of the basal reflexions from the others in the preliminary survey of an unknown material. For a full analysis (except of course in simple cases, such as for instance, a bentonite or china clay deposit) a specialized book, such as the Mineralogical Society's monograph,[42] is necessary.

Dr. J. Thewlis (Atomic Energy Research Establishment, Harwell) gave a description of the methods employed in the identification of radioactive and highly toxic substances. When handling these minute quantities of materials, precautions have often to be employed both to protect the operators and to prevent fogging of the film. In addition, these, and similar materials, are often available only in very small quantities. Satisfactory powder photographs can be obtained from samples weighing considerably less than 1 μg. From the point of view of precautions, different situations arise according to the radiation emitted: α-rays are completely stopped by the capillary tube in which the specimen is enclosed: α-emitters may, however, be extremely toxic since many of them lodge preferentially in the bone where they set up bone sarcoma; the maximum permissible amount of plutonium in the body is less than 1 μg. In order to minimize danger to health the radioactive samples are prepared and sealed into Pyrex capillary specimen tubes in laboratories with special remote handling facilities. These tubes are transferred to the X-ray laboratory only when there is no trace of any external deposit. As a second line of defence the tubes are surrounded by celluloid thimbles or nickel "pill-boxes" to prevent spillage even if the capillary tube should be broken.

β- and γ-emitters are handled in a similar way. Since these radiations have a longer range there is the additional risk of film fogging. This can be minimized by reducing the size of the sample so as to increase the ratio of reflecting to total volume. In addition, the X-ray background is reduced as much as possible by cutting down the beam so that it only just bathes the specimen. A metal screen, such as a nickel foil covering the X-ray film, is often used to reduce the background still further. With suitable precautions samples having activities of a few hundred microcuries can be photographed.

Prof. A. J. C. Wilson (University College, Cardiff) traced the history and development of the A.S.T.M. Index. The use of X-ray powder patterns for identification dates back to 1919 when Hull published one of the first papers in English on powder patterns as such.[43] In 1938 the method became more readily available for routine identification work when Hanawalt, Rinn and Frevel[44] in the U.S.A. published a list of the powder patterns of about 1000 substances and, simultaneously, Boldyrev, Mikheev and Kovalev[45] in the U.S.S.R. produced an index of powder patterns of minerals. The suggestion of using the strongest three lines for identification is due to the former of these workers. In 1941, through the initiative of Prof. W. P. Davey, there was set up a joint sub-committee of the American Society for the Testing of Materials and the Committee on X-ray and Electron Diffraction of the United States National Research Council, which later produced an index containing about 4000 cards. In 1944, at the suggestion of Sir Lawrence Bragg, The (British) Institute of Physics assumed the responsibility of organizing the collection of further data to supplement the original index, this work being directed by Prof. Wilson. In 1949 the whole index was revised, a new and more convenient format of card was adopted, the second and third cards for each substance were abandoned, and numerical and substance indices were provided in book form. The complete index now includes some 5000 substances. The sponsorship of the index has changed hands several times and it is now carried out by the American Crystallographic Association, Prof. Davey continuing as editor. Work is in progress to improve the reliability and accuracy of the index and a compilation of errors is due to be published in 1954. In addition, work of a very high grade is proceeding at the U.S. National Bureau of Standards using carefully purified materials and counter diffractometers. The main part of the future work sponsored by The (British) Institute of Physics may be of the same nature since it is becoming questionable whether further routine photographic work is worth the cost and effort. With properly designed and used counter diffractometers geometrical and absorption aberrations are practically negligible for most substances.

For complex structures, powder photographs are less satisfactory than cell dimensions for identification and the lengths of the axes can be determined very rapidly on a single-crystal camera. Prof. E. G. Cox is investigating the feasibility of preparing an index of substances arranged according to cell dimensions.

Considerable interest was shown by subsequent speakers in the new type of punched index cards of the keysort variety, examples of which were shown in the exhibition. Prof. J. D. Bernal (Birkbeck College, London) pointed out that in some cases the strongest line of a powder pattern is not the best one to use for identification purposes since many classes of related substances have similar prominent spacings owing to the dominant nature of certain common bond lengths. Weak and "lonely" lines in a pattern are often more characteristic of a sample. The longest spacing is also often very useful. He described an index used by Prof. Wyart which contains one card for each line instead of one for each substance. A suggestion was made by him in respect of single-phase patterns which stubbornly resisted identification. Could these not be sent to a central organization of workers in special fields (e.g. alloys) for further scrutiny? He also felt that powder methods would generally not be successful

for identifying complex organic materials. Using single-crystal methods, he had found it possible to identify 10^{-9} g of material from one animal cell.

In the concluding paper entitled "Identification of inorganic and organic compounds by X-ray methods in conjunction with other techniques," MR. R. BROOKS (I.C.I. Ltd., Northwich) confined his observations to the practice in his laboratory in those cases where the use of X-ray methods, and generally those involving a 9 cm diameter powder camera, gives a start to the problem, but does not by itself provide enough evidence for complete identification.

As the first step towards identification, a search is made through the A.S.T.M. index for a match with the three strongest lines of the sample pattern. If this fails, as it often does if the specimen is polyphase, a match is then sought in a laboratory-compiled index of single strongest lines, which of course is much less discriminating. Complications may be present due to the strongest line being compounded from weaker lines of different constituents.

If a stoichiometric analysis is available, or if the origin of the sample is known and a guess can be made of what elements are likely to be present, it is usually possible to eliminate at once most of the matched patterns. An analysis may help in finding minor constituents by subtracting from the analytical data amounts corresponding to the established major constituent or constituents. It is also the best aid in dealing with solid solutions, which can be very difficult to identify unless the basic patterns of end members can be recognized from memory. A subsection of the A.S.T.M. index dealing only with solid solutions would certainly be useful.

Some information on the elements present in a sample can often be obtained by visual examination of line profiles. For example, if the pattern shows lines much thinner than the width of the specimen, this indicates high absorption; it may then be assumed that an element, whose atomic number is just below that of the target element or over about 40, is present to an important extent. More detailed information can be obtained if there is noticeable X-ray fluorescence. For example, if nickel-filtered copper radiation is used, its spectral distribution has the CuKα peak at 1·54 Å and a band whose peak is at about 0·4 Å. The heavy background on the photograph may thus be due to either the characteristic line or to the short wavelength band, exciting fluorescent radiation from the respective elements V_{23}, Cr_{24}, Mn_{25}, Fe_{26}, Co_{27} or Sr_{38} to Ba_{56} and Pb_{82}. It can be determined which group is responsible by taking another photograph using iron and nickel filters together. If the fluorescence is caused by CuKα, the new photograph will show no pattern and only very weak bands at low angles. If it is due to the short wavelength band, again there will be no pattern, but there will be a strong background. Further differentiation can be achieved by using absorption filters placed directedly in front of the film. These can be prepared by impregnating a suitable paper with a suspension of the appropriate finely divided metal oxide in cellulose nitrate solution so that the final density of oxide is about 1 mg/cm^2. The principal wavelength of the fluorescent radiation is easily noted by observing which filters give high absorption. For example, it is readily possible to distinguish between iron and chromium.

There are a number of effects which serve as aids in breaking down the pattern of a complex specimen into the patterns of its constituent phases. It is, of course, necessary to use as high a resolving power of the camera unit as possible to minimize overlapping lines. Broad lines, split lines, spotty lines, asymmetric lines and the previously-mentioned abnormally-thin lines arising from absorption can all be useful in separating patterns. The occurrence of patterns which show distortion or weak or broad lines from phases which are usually good reflectors is a signal to look out for the presence of metastable polymorphs or hydrates. A case arose where the system studied involved only calcium carbonate and water. Apart from the broadened calcite pattern, a single unaccountable line was present, so that a new metastable phase was suspected. Further investigation showed the line to be the strongest in the pattern of calcium carbonate monohydrate.

It is frequently useful to concentrate constituents by physical means and to X-ray the fractions separately. Methods have been used based on particle shape, size, colour, birefringence, ultra-violet fluorescence, density, magnetic properties, etc. If a solvent or reagent is used, it is advisable to check that the residue has not been seriously altered by the treatment.

The occasion often occurs where the material consists of a single crystal which has to be identified and preserved. The full determination of the cell dimensions generally involves too much labour for rapid identification and so a single rotation photograph is taken with a random crystal setting. Interplanar spacings are read off a laboratory-made transparency, or a variation of the method described by Bannister and Hey[46] is used.

Refractive index measurements of transparent grains using the conventional immersion method are often of value. The occasions when morphological observations prove helpful are less frequent. Whilst powder photographs may sometimes give a clear identification, the use of the polarizing microscope is still necessary in providing information on crystal habit variations and in detecting the existence of pseudomorphic secondary products. In one example, the powder patterns indicated calcite, but the polarizing microscope showed that the crystal outlines were characteristic of the normal habit of calcium carbonate hexahydrate.

Organic compounds can sometimes be satisfactorily identified by powder methods, but several curious experiences with substances of moderately high molecular weight have resulted in a diminished confidence in the method. It was possible to point out the existence of an impurity in a body said to be hydroxy–glutaconic–dialdehyde–dianil hydrochloride, but which actually proved to be the monohydrate. However, this material and a similar compound without the hydroxy-group gave indistinguishable X-ray patterns, yet infra-red absorption spectra showed clear differences. Some polyethylenes and oxidized polyethylenes are indistinguishable by X-ray methods alone.

DR. K. W. ANDREWS (United Steel Companies Ltd., Sheffield) spoke about the use of cell dimension/composition curves or interplanar spacing/composition curves of solid solutions. An example where a variety of elements in solid solution would make this approach difficult was also given—the complex carbides of general formula $M_{23}C_6$. With regard to the A.S.T.M. Index, he emphasized the need for having accurate spacings which were either calculated from cell dimensions or corrected by means of an internal standard, in order to detect small solid solution differences.

MR. G. H. COCKETT (Armament Research Establishment, Woolwich) urged the advantages in using a curved lithium fluoride crystal monochromator with a 4½ in. diameter transmission focusing camera in identification problems. Using this method the exposure time is appreciably less than that required in a 9 cm powder camera using filtered radiation, the exposure ratio between lithium fluoride and quartz focusing monochromators being about 10 : 1. The technique

is particularly useful in identifying a mixture of compounds giving weak diffuse diffraction lines. In addition Mr. Cockett uses flat and curved lithium fluoride crystals in fluorescence analysis cameras. With these cameras it is possible to distinguish between iron and chromium in 2 min: the identification of mixed carbides discussed by Dr. Andrews could readily be carried out by fluorescence analysis.

In reply to a question concerning the determination of the amount of manganese in an aluminium-manganese alloy, DR. W. A. WOOSTER (University of Cambridge) mentioned the method involving the determination of absorption coefficients using a Weissenberg goniometer.[47] He pointed out that a flake of thickness $\frac{1}{2}$ mm or less, containing small quantities of such elements that show strong absorption for the usual X-radiations, is suitable for use with this method. Secondly, he commented on the use of punched cards for classification of substances and illustrated the principle by suggesting that, if cards punched round all sides were used, one side might be devoted to the spacing, another to the first letter of the formula, a third side to the second letter and the fourth, perhaps, to the spacing of "lonely" lines.

DR. H. S. SIMON (Birkbeck College, London) mentioned that in examining very finely-divided "fly-ash" from power station boilers, he had found it all completely isotropic under the petrological microscope, in spite of the fact that it showed many sharp X-ray lines which certainly did not derive from substances of cubic symmetry. Gravimetric and magnetic methods of separating the various phases were not successful as the magnetic constituents were attached to the bubbly glass particles.

An example of the use of electron diffraction methods in identification problems was given from the floor; by means of a three-stage electron microscope it had been possible to obtain diffraction patterns from single clay particles less than 1 μ in diameter.

In addition to the sessions reported there was a business session devoted to a lively and profitable discussion of the activities of the Equipment Sub-committee of the Group. A stimulating evening discourse entitled "Some thoughts on the future of natural science, with illustrations from the growth of X-ray diffraction work in the United States,"[48] was delivered by DR. R. W. G. WYCKOFF (U.S. Embassy, London). A valuable feature of the conference was the unusually extensive exhibition of crystallographic equipment; commercially available apparatus and examples of recent research developments were shown.

U. W. ARNDT J. B. NELSON

REFERENCES

(1) GURNEY, R. W., and MOTT, N. F. *Proc. Roy. Soc. A*, **164**, p. 151 (1938).
(2) SEEMANN, H. E. *Rev. Sci. Instrum.*, **21**, p. 314 (1950).
(3) CLAASSEN, H. H., and BEU, K. E. Private communication.
(4) CORNEY, G. M., and SPLETTSTOSSER, H. R. *Non-Destructive Testing*, **10**, p. 29 (1951).
(5) RONNEBECK, H. R. *J. Sci. Instrum.*, **20**, p. 154 (1943).
(6) TAYLOR, A. *J. Sci. Instrum.*, **28**, p. 200 (1951).
(7) DOBSON, G. M. B. *Proc. Roy. Soc. A*, **104**, p. 248 (1923).
(8) ASTBURY, W. T. (i) *Proc. Roy. Soc. (A)*, **115**, p. 640 (1927). (ii) *Proc. Roy. Soc. (A)*, **123**, p. 575 (1929).
(9) ROBINSON, B. W. *J. Sci. Instrum.*, **10**, p. 233 (1933).
(10) ROBERTSON, J. M., and DAWTON, R. H. V. M. *J. Sci. Instrum.*, **18**, p. 126 (1941).
(11) WIEBENGA, E. H. *Rev. Trav. Chim. Pays-Bas*, **66**, p. 746 (1947).
(12) COCHRAN, W. *Acta Cryst.*, **3**, p. 268 (1950).
(13) BRENTANO, J. *Z. Phys.*, **70**, p. 74 (1931).
(14) BRENTANO, J., BAXTER, A., and COTTON, F. W. *Phil. Mag.*, **17**, p. 370 (1934).
(15) KLUG, H. P. *Analyt. Chem.*, **25**, p. 704, 1953.
(16) HALL, W. H., ARNDT, U. W., and SMITH, R. A. *Proc. Phys. Soc. A*, **62**, p. 631 (1949).
(17) COYLE, R. A., HALE, K. F., and WAINRIGHT, C. *J. Sci. Instrum.*, **30**, p. 151 (1953).
(18) WILLIAMSON, G. K., and SMALLMANN, R. E. *J. Sci. Instrum.*, **30**, p. 341 (1953).
(19) GILLAM, E., and COLE, D. G. *J. Sci. Instrum.*, **30**, p. 429 (1953).
(20) GORDON, R. L. *J. Sci. Instrum.*, **30**, p. 431 (1953).
(21) COOKE-YARBOROUGH, E. H., FLORIDA, C. D., and DAVEY, C. N. *J. Sci. Instrum.*, **26**, p. 124 (1949).
(22) MUEHLHAUSE, C. O., and FRIEDMAN, H. *Rev. Sci. Instrum.*, **17**, p. 506 (1946).
(23) PARRISH, W., and HAMACHER, E. A. (i) *Science*, **110**, p. 368, 1949. (ii) *Trans. Instruments and Measurement Conference, Stockholm*, p. 95, 1952.
(24) LANG, A. R. *Nature*, **168**, p. 907 (1951).
(25) ARNDT, U. W., and RILEY, D. P. *Proc. Phys. Soc. A*, **65**, p. 74 (1952).
(26) LANG, A. R. *Proc. Phys. Soc. A*, **65**, p. 372 (1952).
(27) ARNDT, U. W. *Atomics*, **4**, p. 217 (1953).
(28) ARNDT, U. W., COATES, W. A., and CRATHORN, A. R. *Proc. Phys. Soc. B*, **67**, p. 357 (1954).
(29) BIRKS, J. B. *Scintillation Counters* (London: Pergamon Press, Ltd., 1953).
(30) BACON, R. C., and POLLARD, J. R. *Electronic Engng*, **22**, p. 173 (1950).
(31) HOUGH, G. H., and RIDLER, D. S. *Electrical Communication*, **27**, p. 214 (1950).
(32) BLIFFORD, I. H., ARNOLD, R. G., and FRIEDMAN, H. *Electronics*, **22**, p. 110 (1949).
(33) KANDIAH, K. *Proc. Instn Elect. Engrs*, **101**, p. 227 (1954).
(34) BACON, G. E., and LONSDALE, K. *Rep. Progr. Phys.*, **16**, p. 1 (1953).
(35) BACON, G. E. *Proc. Roy. Soc. A*, **209**, p. 397 (1951).
(36) BACON, G. E., and LOWDE, R. D. *Acta Cryst.*, **1**, p. 303 (1948).
(37) BACON, G. E., and THEWLIS, J. *Proc. Roy. Soc. A*, **196**, p. 50 (1950).
(38) BACON, G. E., SMITH, J. A. G., and WHITEHEAD, C. D. *J. Sci. Instrum.*, **27**, p. 330 (1950).
(39) BUNN, C. W. *Chemical Crystallography*, p. 379 (Oxford: Clarendon Press, 1945).
(40) ITO, T. *Nature*, **164**, p. 755 (1949).
(41) HODGMAN, C. D., WEAST, R. C., and WALLACE, C. W. *Handbook of Chemistry and Physics*, 35th Ed. (Cleveland: Chemical Rubber Publishing Co., 1953).
(42) *X-ray Identification and Crystal Structures of Clay Minerals*. Editor G. W. Brindley (London: The Mineralogical Society, 1951).
(43) HULL, A. W. *Phys. Rev.*, **10**, p. 661 (1917).
(44) HANAWALT, J. D., RINN, H. W., and FREVEL, L. K. *Industr. Engng Chem.* (Anal. Ed.), **10**, p. 457 (1938).
(45) BOLDYREV, A. K., MIKHEEV, V. I., and KOVALEV, G. A. *Zapiski Gornogo Instituta*, **11**, (2), p. 1 (1938). There are some later supplements.
(46) BANNISTER, F. A., and HEY, M. *Miner. Mag.*, **24**, p. 49 (1935).
(47) WOOSTER, W. A. *Miner. Mag.*, **29**, p. 427 (1950).
(48) WYCKOFF, R. W. G. *Brit. J. Appl. Phys.*, **5**, p. 199 (1954).

Determination of thermal conductivities at high temperatures

By Sir K. S. Krishnan, F.Inst.P., F.R.S., and S. C. Jain, M.Sc., National Physical Laboratory of India, New Delhi

[*Paper first received* 21 *April, and in final form* 12 *July*, 1954]

The usual electrical methods of determining thermal conductivities either cannot be used at high temperatures or their use is restricted to substances like carbon or graphite which satisfy certain special requirements. It is shown that the observed temperature distribution along a metal filament electrically heated in a vacuum can be used for the determination of the thermal conductivity at high temperatures. The temperature distribution near the centre of a short filament, which is known to be parabolic, or the distribution in regions slightly removed from the centre of a long filament, which is known to be logarithmic, are particularly suitable for this purpose. Measurements are reported on the thermal conductivity of platinum in the temperature region 1300 to 1800° K made in this manner.

THE DIFFERENTIAL EQUATION DEFINING THE STEADY STATE

Consider a filament electrically heated in a vacuum, so that the transfer of heat by convection in the surrounding atmosphere is avoided. The heat generated per unit length of the filament per second is $I^2\rho/\omega$, where I is the heating current, ρ is the specific resistance, and ω is the area of cross-section of the filament. The rate of loss of energy per unit length by radiation is $p\epsilon\sigma(T^4 - T_0^4)$, where p is the periphery of the cross-section, ϵ is the total emissivity from the surface, as distinguished from the spectral emissivity ϵ_λ, and σ is Stefan's constant of radiation. Owing to the conduction of heat towards the ends, the energy abstracted per second per unit length is $-\kappa\omega d^2T/dx^2$, where κ is the thermal conductivity. Hence the steady state is determined by the well-known differential equation[1]

$$\kappa\omega d^2T/dx^2 - p\epsilon\sigma(T^4 - T_0^4) + I^2\rho/\omega = 0 \qquad (1)$$

Now the existence of a temperature gradient along the filament is a consequence of the finite thermal conductivity of the filament. Hence if we had an analytical solution of equation (1), we might utilize the observed distribution of temperature along the filament to determine experimentally the thermal conductivity. But an analytical solution of the differential equation in the general case was not available till recently. Hence some of the earlier experimental methods for the determination of κ, as for example those of Kohlrausch,[2] and of Callendar,[2] were so designed as to reduce equation (1) to a readily integrable form, namely to

$$d^2T/dx^2 = f \qquad (2)$$

where f is an explicit function of κ. Obviously the solution is

$$T = \tfrac{1}{2}fx^2 + Bx + C \qquad (3)$$

in which the constants B and C may be determined from the boundary conditions. The actual techniques used in these methods for securing equation (3), however, restrict the applicability of these methods to temperatures in the neighbourhood of the room temperature.

On the other hand Worthing and Halliday[3] have adopted a graphical method for the determination of the thermal conductivity from the observed distribution of temperature in heated filaments, which in essence is equivalent to solving the differential equation (1) graphically. Knowing all the physical constants involved in equation (1) except κ, κ then becomes known. Actually they integrate graphically the total heat generated in a selected finite length of the filament, and similarly also the total heat radiated out from this length. The difference between the two gives $\kappa\omega(G_1 - G_2)$, where G_1 and G_2 are the temperature gradients at the colder and the hotter ends respectively of the length selected. By making one end of the selected length coincide with the centre of the filament, G_2 can be made zero.

This method of determining κ is obviously applicable to high temperatures also.

In some recent papers[4,5,6] we have investigated in detail the analytical solution of equation (1). In the case of a long filament it is possible to obtain a general solution of equation (1), and in the case of shorter filaments too a solution can be obtained in terms of a power series which is rapidly convergent over most parts of the filament. It is the main purpose of this paper to show that the observed temperature distribution along such filaments, which can be readily measured with an optical pyrometer, can be made the basis of a convenient and direct method for determining thermal conductivities at high temperatures. This is of interest experimentally. The only other method that is available now, namely, that of Angell,[7] is based on the measurement of the difference in temperature between the inner and the outer surfaces in the middle portion of a long tube electrically heated in a vacuum. This method has been used by Powell and Schofield[8] to determine the thermal conductivity of carbon and of Acheson graphite; in which the walls of the tube can be made thick enough to secure a considerable difference in temperature between the inner and the outer surfaces, and at the same time the electrical resistance of the tube remains sufficiently large to permit our heating it to high temperatures without using inconveniently large currents. This cannot, however, be done with good metals.

It is significant that the only high temperature data for thermal conductivities available at present are those for tungsten, tantalum and carbon obtained by Worthing's graphical method, and those for carbon and graphite obtained by Angell's method.

THE METHODS OF KOHLRAUSCH AND CALLENDAR AS DEVICES FOR SIMPLIFYING THE TEMPERATURE DISTRIBUTION

In Kohlrausch's method, which has been developed experimentally by Jaeger and Diesselhorst,[2] the loss by radiation is practically eliminated by surrounding the conductor with a suitable wadding of insulating material. The differential equation (1) now reduces to

$$d^2T/dx^2 + I^2\rho/(\kappa\omega^2) = 0 \qquad (4)$$

which is readily solved.

In Kohlrausch's method the distances x are not measured directly, but in terms of the corresponding potential drops

$$v = (I\rho/\omega)x \tag{5}$$

Since there is no loss by radiation, and therefore no gradient of temperature *across* the length, one may use, as these authors do, a rod or a bar of large cross-section, instead of a filament.

On the other hand, in Callendar's method a low heating current is used so as to make $T - T_0$ sufficiently small over the whole length of the filament that the radiation loss may be regarded as proportional to $T - T_0$ instead of to $T^4 - T_0^4$, that is, the radiation cooling follows Newton's law. The temperature range being now small, the variation of the specific resistance with temperature may be taken to be linear. Hence the resistance at every point may be split into a constant term characteristic of the temperature T_0, and an extra term proportional to $T - T_0$. Consider now the heating due to the latter part of the resistance, which will be proportional to I^2 and to $T - T_0$. By suitably adjusting the heating current I, this part of the heat generated can be made to balance the loss due to radiation, at every point of the filament, since both are proportional to $T - T_0$. When this condition is secured, the energy that is abstracted per unit length in the process of conduction becomes constant, i.e. the same at all points along the filament. This is essentially the condition for the temperature distribution to be parabolic. In the actual measurements the increase in resistance due to the heating, which can be readily calculated in terms of the parabolic distribution, is measured, rather than the distribution of temperature as such.

Both these methods, by the limitations imposed by the special techniques involved in securing the parabolic distribution, are applicable only to temperatures in the neighbourhood of the room temperature.

We now proceed to show that independently of any such special devices the temperature distribution near the centre of any finite filament is naturally parabolic, and can be utilized in the same manner as in Kohlrausch's and Callendar's special arrangements, to determine κ. Unlike these special arrangements, this method is applicable to high temperatures too.

ACTUAL DISTRIBUTION NEAR THE CENTRE SHOWN TO BE PARABOLIC AND UTILIZED TO DETERMINE κ

Let T_l be the temperature at the centre of the filament, and T_m the value to which T_l tends as $2l$, the length of the filament, is increased indefinitely, keeping the heating current constant. Obviously

$$p\sigma\epsilon(T_m^4 - T_0^4) = I^2\rho/\omega \tag{6}$$

Returning to equation (1), and eliminating T_0 with the help of equation (6), one obtains

$$\frac{d^2T}{dx^2} + \frac{p\epsilon\sigma}{\kappa\omega}(T_m^4 - T^4) = 0 \tag{7}$$

which in the case of a finite rod for which $T_l \neq T_m$ can be written in the form

$$\frac{d^2T}{dx^2} + \frac{p\epsilon\sigma}{\kappa\omega}(T_m^4 - T_l^4) + \frac{p\epsilon\sigma}{\kappa\omega}(T_l^4 - T^4) = 0 \tag{8}$$

If $T_m - T_l$ is considerable, as will be the case when the filament is not long, one may confine attention to a region close to the centre, where the temperature-dependent term in equation (8), namely, the third term on the left-hand side, becomes negligible in comparison with the second term. Equation (8) then reduces to

$$d^2T/dq^2 + f_1(\kappa) = 0 \tag{9}$$

where

$$f_1(\kappa) = \frac{p\epsilon\sigma}{\kappa\omega}(T_m^4 - T_l^4) \tag{10}$$

and $q = l - x$ is now the distance measured from the centre. Using the boundary conditions that when $q = 0$, $T = T_l$ and $dT/dq = 0$, one obtains

$$t = T_l - T = \tfrac{1}{2}f_1q^2 \tag{11}$$

Thus from the observed temperature distribution in this region it is possible to determine f_1 and then κ in the same manner as before. Since the largest deviation of the temperature in the parabolic region from T_l, that is, the highest value of t involved in this region, is small, the quantity $p\epsilon/(\kappa\omega)$ appearing in equation (10) for f_1 may be assigned the value appropriate to the temperature T_l.

κ FROM THE LOGARITHMIC REGION

Consider a long filament in which $T_l \simeq T_m$. In this case, if one is not too close to the centre, that is, if T is not too close to T_l, the second term on the left-hand side of equation (8) may be neglected in comparison with the third. Consider a region sufficiently removed from the centre, but where $T_m - T = \Delta$ is still small in comparison with T_m. Equation (8) then takes the simple form

$$d^2\Delta/dx^2 = A\Delta \tag{12}$$

where

$$A = \frac{4p\epsilon\sigma T_m^3}{\kappa\omega} \tag{13}$$

Neglecting for the present the temperature variation of the physical quantities involved, one then obtains the well-known solution

$$x\sqrt{A} = D - \ln\Delta \tag{14}$$

in which $1/\sqrt{A}$ is obviously the distance over which Δ changes by a factor e. Hence A can be determined experimentally, and κ can be evaluated therefrom.

We should mention here that the temperature distribution along a thin-walled tube electrically heated in a vacuum is somewhat similar to that along a filament. When the tube is long, the temperature distribution in regions slightly removed from the centre is given by equation (14), in which A has the same significance as in a filament, except that the cross-sectional area ω occurring in the expression for A is now the area of the annular ring. A part of A, namely $\epsilon\sigma/\kappa$, occurs also in the expression for the difference in temperature between the inner and the outer surfaces of the tube. In determining this difference in temperature Prescott and Hincke[9] have used the observed temperature distribution in the logarithmic region, in order to eliminate $\epsilon\sigma/\kappa$ from the expression for this difference. Obviously the experimental data for the temperature distribution in the logarithmic region could have been used to determine κ also.

Coming back to the filament, we should mention also the following results which were derived in our earlier papers, and which are relevant to our present purpose.

(1) Though the condition under which equation (14) has been derived, namely that Δ be small in comparison with T_m, is a sufficient one, it is not necessary, the necessary condition being that $\Delta/\ln\Delta$ be small in comparison with T_m, which

obviously holds over a much wider range than the former condition.

(2) The region near the centre that is excluded corresponds to a total range of temperature equal to a few times $T_m - T_l$, and hence experimentally trivial.

(3) Now in formulating equation (12) we have neglected for convenience the temperature variations of the different physical quantities involved, namely $\kappa\omega$, ρ/ω and $p\epsilon$. The effect of the temperature coefficients of these quantities on equation (12) is essentially different from that in the parabolic region. As we shall see presently, even when attention is confined to a *narrow* region in the logarithmic range, the effect of the temperature coefficients will be appreciable, unlike in the parabolic range. This is due to the following circumstance. In the differential equation (9), which defines the temperature distribution in the parabolic range, the effect of the temperature coefficients is in the form of a contribution to the term proportional to t, which in any case may be neglected since t is chosen small. In other words equation (9) will hold even when the temperature coefficients are considerable, provided that the value of $p\epsilon/(\kappa\omega)$ occurring in the expression for f_1 is made to refer to the temperature T_l.

On the other hand in the logarithmic region the effect of the temperature variation of ρ/ω and $p\epsilon$ is to introduce an extra small term which varies linearly with temperature, which now cannot be neglected. Taking the values of $\kappa\omega$, ρ/ω and $p\epsilon$ to refer to the temperature T_m, the effect of the finite temperature coefficients of these quantities—the temperature coefficient of $\kappa\omega$ has very little effect in this region—is to introduce in equation (12) a small extra term proportional to Δ. In other words the expression for A in equations (12) and (13) will include a multiplying factor a_1 of the order of unity, whose deviation from unity is determined by the temperature coefficients of ρ/ω and $p\epsilon$, and by T_m.

The calculation of a_1 was given in a recent paper,[4] in which, however, the coefficient of thermal expansion, which normally is much smaller than the coefficients of κ, ρ and ϵ, was neglected. The expression for a_1 obtained in the paper can, however, be made to include thermal expansion also by replacing the temperature coefficients of κ, ρ and ϵ given in that paper by the coefficients of $\kappa\omega$, ρ/ω and $p\epsilon$ respectively.

We shall merely mention here the final result, namely, that for platinum, with which we shall be concerned in the present paper, a_1 varies from 0·97 for $T_m = 1800°$ K to 1·00 for $T_m = 1300°$ K.

TEMPERATURE DISTRIBUTION IN OTHER REGIONS ALSO UTILIZED TO GIVE κ

In the case of a long filament one can also obtain a general solution of equation (1) applicable over the whole length of the filament. Neglecting provisionally the temperature coefficients, the solution is

$$\Delta \exp [f(\Delta)] = \Delta_0 \exp [f(\Delta_0)]\{\exp(-x\sqrt{A}) + \exp[-(2l-x)\sqrt{A}]\} \quad (15)$$

where

$$\Delta_0 = T_m - T_0 \quad (16)$$

T_0 being the temperature of the point from which x is measured, and

$$f(\Delta) = \tfrac{1}{2}\Delta/T_m + \tfrac{1}{6}\Delta^2/T_m^2 - \ldots \quad (17)$$

Now in a long filament the region near the centre where $\exp[-(2l-x)\sqrt{A}]$ is significant, is not of experimental interest since the total variation of temperature in this region is a few times $T_m - T_l$, which is negligible. Outside this region, where $\exp[-(2l-x)\sqrt{A}]$ is negligible in comparison with $\exp(-x\sqrt{A})$, equation (15) reduces to the simple form

$$\Delta \exp [f(\Delta)] = \Delta_0 \exp [f(\Delta_0)] \exp(-x/A) \quad (18)$$

Now expressing all distances in terms of $1/\sqrt{A}$ and all temperatures in terms of T_m, that is, using the reduced co-ordinates $X = x\sqrt{A}$ and $\tau = T/T_m$, equation (18) may be seen to reduce to

$$(1-\tau) \exp [\phi(\tau)] = (1-\tau_0) \exp [\phi(\tau_0)] \exp(-X) \quad (19)$$

where $\phi(\tau) = f(\Delta)$. Obviously equation (19) is a very general expression, which implies that a *single* $\tau - X$ curve will represent the temperature distribution in all *long* filaments, independently of the nature of the filament, its length, its cross-section, or even the heating current used, provided that the reduced temperature τ_0 of the point from which the distances are measured is chosen the same in all the plots.

Hence, knowing T_m in any actual experiment, one can easily plot τ against x. This curve may be compared with the theoretical plot of τ against X, taking τ_0 from which x is measured, as the origin of distance in the $\tau - X$ curve too. The only difference between the two curves should be in the scales representing the abscissae. In other words it should be possible to transform the experimental $\tau - x$ curve into the theoretical $\tau - X$ curve by increasing the abscissae of the former by a *constant* factor F, that is, by a factor which is independent of x, and which can be experimentally determined. This factor should obviously be equal to $\sqrt{A}$, from which κ can be determined.

Here again we have neglected the temperature coefficients of the different quantities. Taking them into account would introduce a factor a_1 in the expression for A, as in the logarithmic region, and a factor a_2 in the expression for $\phi(\tau)$. For platinum a_2 varies from 1·4 for $T_m = 1800°$ K, to 1·8 for $T_m = 1300°$ K.

Confining now attention to the top of the region that we are considering, in which Δ is small enough to keep $\exp \phi(\tau)$ close to unity, equation (19) will obviously reduce to the logarithmic formula, as it should.

In reduced co-ordinates the temperature distribution in the logarithmic region is given by

$$\ln(1-\tau) = \ln(1-\tau_0) + \phi(\tau_0) - X \quad (20)$$

Hence, if the reduced temperature τ_0 from which distance measurements are made is taken to be the same in all the measurements, the plot of $\ln(1-\tau)$ against the distance will be a series of straight lines, all of them passing through the point τ_0, and having different slopes. The slopes give directly $\sqrt{A}$, and hence κ, taking into account the temperature coefficients of ρ/ω and $p\epsilon$.

When one moves out of the logarithmic region $\phi(\tau)$ grows fairly rapidly and the effect of the temperature coefficients on $\phi(\tau)$ also becomes significant. Now $\phi(\tau)$ involves in addition to the temperature coefficients of ρ/ω and $p\epsilon$ which are known, the temperature coefficient of $\kappa\omega$ too. Hence in the same manner in which temperature distribution in a long filament in the logarithmic range can be used to give κ, the distribution outside the logarithmic range can be made to yield both κ and its temperature coefficient α.

In practice, however, it is found more convenient to determine κ at different temperatures from the logarithmic range, by suitably varying the heating currents and varying the temperatures T_m to which κ refers. This has been done in the measurements described below.

EXPERIMENTAL

It has been shown that the thermal conductivity can be determined from the temperature distribution either in the parabolic region near the centre of any finite filament, preferably a short one, or in the logarithmic or any other region of a long filament. We have made measurements with a platinum wire, of spectroscopic purity (obtained from Johnson and Matthey and Co. Ltd.). The main part of the experiment is the determination of the temperature at different points along the heated wire. This was done with an optical pyrometer and for this purpose one needs to know the spectral emissivity ϵ_λ for the spectral region utilized in the pyrometer, namely the neighbourhood of $0{\cdot}655\ \mu$, at different temperatures.

This was determined in the following manner. In a recent paper we described some measurements on the spectral emissivity of Acheson graphite.[10] Using a thin-walled, long tube of graphite, electrically heated in a vacuum, the temperature of the inner surface of the tube near its centre is readily obtained by boring a small hole in the wall, and by measuring directly with the pyrometer the temperature of the cavity as viewed through the hole. The small difference in temperature between the inner and the outer surfaces is estimated easily from any rough value of the thermal conductivity. Thus the temperature of the outer surface near the centre is known. Knowing also the brightness temperature of the surface as measured with the optical pyrometer, the spectral emissivity ϵ_λ is obtained directly. By suitably adjusting the heating current, the spectral emissivity can be obtained at different known temperatures. Conversely, knowing ϵ_λ, the actual temperature of the surface corresponding to any observed brightness temperature can also be readily obtained.

For our present purpose, namely, the determination of the spectral emissivity of platinum at different temperatures, we used one of these thin-walled long graphite tubes, and wound round it in a thin groove cut round the middle of the tube, a single ring of platinum wire. The presence of either the groove or the platinum wire in the middle does not affect the temperature distribution. Knowing the actual temperature of the wire, which is that of the graphite surface adjoining it, and the brightness temperature of the wire as obtained with the optical pyrometer, its spectral emissivity is known. By suitably varying the heating current, values of ϵ_λ can be obtained at different temperatures. They are plotted in Fig. 1.

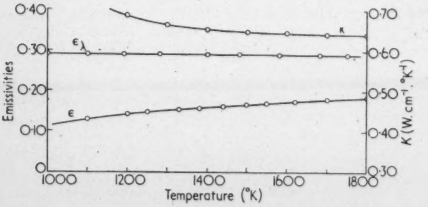


Fig. 1. The temperature variation of the thermal conductivity of platinum and of the spectral and the total emissivities

The first series of measurements that we made on the spectral emissivity of platinum was with a foil rolled into a cylinder, which just surrounded the graphite tube without touching it. It was fixed to the supporting block at one end. The external diameter of the graphite tube was 8 mm, and the clearance space between the graphite tube and the tube of platinum foil was about $\frac{1}{2}$ mm. The tube of platinum foil extended over nearly the whole length of the graphite tube, but not completely. This was to avoid passage of electric current through the platinum foil. Hence the temperature of the foil, particularly near the centre, where it is sensibly constant, will be just that of the graphite tube underneath, and that of the annular space between the two tubes. A small hole of about 1 mm diameter punctured in the platinum foil, and also in the wall of the graphite tube just underneath it, enabled the temperature in the cavity near the middle of the graphite tube to be measured directly.

Knowing this temperature, and the small difference between the inner and the outer surfaces of the graphite tube, the temperature of the platinum surface adjoining the hole is known. From this temperature and the brightness temperature of the platinum surface as measured with the pyrometer, the spectral emissivity of platinum is readily determined.

The spectral emissivity data for platinum obtained in this manner were found to agree closely with the measurements reported, which were made with a ring of platinum wire wound round the central portion of the graphite tube.

The available data for the spectral emissivity of platinum obtained by different authors vary widely, and the differences have sometimes been attributed to the probable differences of the surfaces of the specimens used. The close agreement between our measurements with the surface of the platinum foil and of the wire does not, however, support this view. Even so, since the specimen with which the thermal conductivity measurements were made, was in the form of a wire, we considered the measurements on ϵ_λ made with the same specimen to be more appropriate for our present purpose than the measurements made with the foil. The results of the latter measurements, which are not incorporated in the graph, are given in the following table.

T° K	1120	1250	1400	1525	1600	1690	1750
ϵ_λ	0·28	0·30	0·29	0·28	0·29	0·29	0·30

The single coil of the platinum wire wound in a small groove round the middle of the graphite tube involves negligible change in the heat capacity, and is therefore not likely to disturb the temperature distribution sensibly. The agreement between the values of ϵ_λ obtained by the two methods, supports this conclusion. Observations made with the optical pyrometer also confirmed that the temperature of the graphite surface in the close neighbourhood of the ring of platinum wire did not show any detectable variation.

The experimental error in the measurement of ϵ_λ is estimated to be less than $\pm 0{\cdot}02$.

Referring again to the temperature measurements, two sets of temperature measurements were made on platinum wires of 0·5 mm diameter (s.w.g. 25) heated in a vacuum by an alternating current. In the first set the wire was a medium long one, and the temperature measurements were made in the parabolic region near the centre. In Fig. 2 are plotted the values of t against q^2. In one series of measurements T_m and T_l were 1600 and 1400° K respectively, while in the other series 1500 and 1200° K respectively. The plot can be seen to be a straight line for small values of q^2, and its slope according to equation (11) gives $\frac{1}{2} f_1$, from which κ can be calculated. It can also be seen that t increases slightly more rapidly than in proportion to q^2, since the last one or two points are slightly above the straight line. This is to be

expected theoretically. But the straight line portion is long enough to enable the slope to be determined accurately. The values of κ for 1200 and 1400° K plotted in Fig. 1 were obtained in this manner.

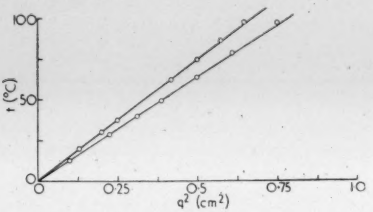


Fig. 2. The parabolic variation of temperature near the centre of the filament

In the second set measurements were made with a *long* platinum wire in the logarithmic region, using different heating currents. In Fig. 3 are plotted the values of $\ln(\Delta/T_m)$ against the distance. The plot is a straight line over the whole of the temperature region included in the curves. The slope gives $\sqrt{A}$, from which again κ can be calculated.

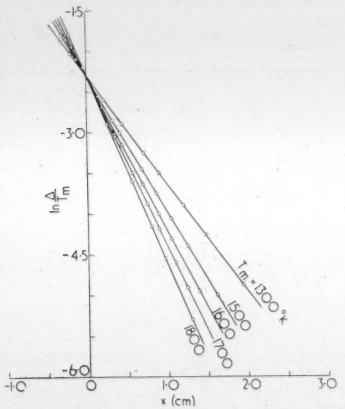


Fig. 3. Temperature variation in the logarithmic region

In order to obtain κ from f_1 or from A one needs to know also the total emissivity ϵ, as distinguished from the spectral emissivity. All that one needs for this purpose is the potential gradient g in the centre of a long wire heated by different currents, since

$$p\epsilon\sigma(T_m^4 - T_0^4) = Ig \qquad (21)$$

ϵ in equation (21) is the value of the total emissivity at temperature T_m.

The potential gradient g is easily measured with the help of two probes. The values of ϵ thus obtained are also plotted in Fig. 1. It is very significant that, whereas the spectral emissivity is almost independent of temperature, the total emissivity drops with the decrease of temperature.

DATA FOR THERMAL CONDUCTIVITY

The values of κ obtained from the parabolic and the logarithmic regions are also plotted in Fig. 1. The values for 1200 and 1400° K were obtained from the former region, and the remaining values from the latter. It will be seen that both the sets plot smoothly on a single curve.

It will also be seen that κ tends to a constant value at high temperatures, as should be expected at temperatures much above the Debye temperature.

THE LORENZ RATIO FOR PLATINUM

In the course of the determination of the total emissivity described in a previous section, one obtains incidentally from the potential gradient and the current, the electrical resistivity at different temperatures. Since we now have data for the thermal conductivities also, we can obtain directly the values of the Lorenz ratio $\kappa\rho/T$ at different temperatures. Since it is an important physical quality, for which at present data are not available for platinum at high temperatures, we have plotted in Fig. 4 this ratio.

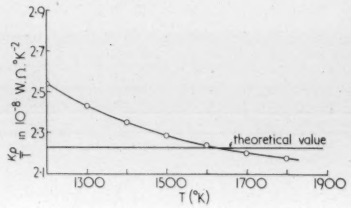


Fig. 4. Lorenz ratio for platinum

It will be seen from Fig. 4 that the Lorenz ratio for platinum decreases slightly with the increase of temperature, and remains close to the theoretical value of $2{\cdot}23 \times 10^{-8}$ W Ω °K^{-2}.

REFERENCES

(1) Carslaw, H. S., and Jaeger, J. C. *Conduction of heat in solids*, p. 135 (London: Oxford University Press, 1947).

(2) For a good account of the electrical methods of determining thermal conductivities see Glazebrook's *Dictionary of Applied Physics*, Vol. I, p. 446 (London: Macmillan and Co. Ltd., 1922).

(3) Worthing, A. G., and Halliday, D. *Heat*, p. 171 (New York: John Wiley and Sons, Inc., 1948).

(4) Jain, S. C., and Krishnan, K. S. *Proc. Roy. Soc. A*, **222**, p. 167 (1954).

(5) Jain, S. C., and Krishnan, K. S. *Proc. Roy. Soc. A*, **225**, p. 1 (1954).

(6) Jain, S. C., and Krishnan, K. S. *Proc. Roy. Soc. A*, **225**, p. 7 (1954).

(7) Angell, M. F. *Phys. Rev.*, **33**, p. 421 (1911).

(8) Powell, R. W., and Schofield, F. H. *Proc. Phys. Soc.*, **51**, p. 153 (1939).

(9) Prescott, C. H., and Hincke, W. B. *Phys. Rev.*, **31**, p. 130 (1928).

(10) Jain, S. C., and Krishnan, K. S. *Proc. Roy. Soc. A*, **213**, p. 143 (1952).

The two-dimensional magnetic or electric field of a single isolated pole-piece

By N. H. Langton, Ph.D., A.Inst.P., and N. Davy, D.Sc., The University of Nottingham

[*Paper first received* 21 *April, and in final form* 11 *June*, 1954]

The two-dimensional magnetic or electric field of a single isolated pole-piece consisting of a thick parallel plate terminated by a concave semi-circular cylinder is investigated theoretically. The method involves the use of conformal transformations and elliptic functions. The variation of the field strength along the edge of the pole-piece and along the external axis of symmetry is calculated and shown graphically.

The field of a single isolated pole-piece consisting of a single thick semi-infinite parallel plate terminated by a convex semi-circular cylinder has already been investigated,[1] and also the field inside the corresponding slit.[2] The other case of practical importance is the two-dimensional field of a single thick semi-infinite plate with parallel sides terminated by a concave semi-circular cylinder, whose cross-section is shown in Fig. 1. The field inside the same-shaped slit has already been investigated,[3] and in this present paper the field external to the body is obtained. The diagram of Fig. 1 may be regarded as a cross-section of a thick pole-piece of magnetized ferromagnetic material extending to infinity in one direction, or else as the outline of a very thin plate. Since the mathematical calculations for two-dimensional systems are the same whether the field is magnetic or electric, the pole-piece may also be considered as an electric conductor with a surface charge. Since the method of solution used here is the same as that used in the papers referred to above, these references should be consulted for full details of the method.

This problem may also have some application in geophysical prospecting, since the cross-section investigated is similar to that taken up by mineral dykes.

If the solution refers to a magnetized pole-piece, then the conditions of the problem are that the magnetic material must have a large and constant permeability, and must be in an unsaturated state. The direction of magnetization of the specimen is along the vertical axis of Fig. 1. The lines of magnetic induction are therefore parallel to the straight sides except near the surface where, because the surface is an equipotential, they curve so as to intersect it at right angles. This is indicated by the dotted lines in Fig. 1.

THE GEOMETRICAL TRANSFORMATION

To perform the geometrical transformation, we note that the edge *ABDEF* of the conductor shown in Fig. 1 can be

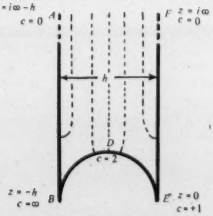


Fig. 1. The z-plane

regarded as a curvilinear triangle in the z-plane, whose external angles are all 2π radians. To transform this figure so that *ABDEF* becomes the real axis of a c-plane, where $c = k^2$ and k is the elliptic modulus, we require two independent solutions of the hypergeometric equation

$$d^2z/dc^2 + \{[\gamma - (\alpha + \beta + 1).c]/c(1 - c)\}.dz/dc - \alpha\beta z/c(1 - c) = 0 \quad (1)$$

With the corner angles 2π we obtain two permissible sets of values for the elements α, β and γ, which are

$$\left.\begin{array}{llll} \alpha = -1/2 & \beta = +3/2 & \gamma = +3 \\ \alpha = +7/2 & \beta = +3/2 & \gamma = +3 \end{array}\right\} \quad (2)$$

Forsyth[4] gives as a solution of equation (1)

$$z = A.\int_0^1 v^{\beta-1}.(1 - v)^{\gamma-\beta-1}.(1 - cv)^{-\alpha}.dv \quad (3)$$

where A is a constant and v is any variable. This solution may be written $z = F(\alpha, \beta, \gamma, c)$. Using the first set of values for the elements we obtain $z = F(-1/2, +3/2, +3, c)$ which gives the integral

$$z_1 = A.\int_0^1 v^{1/2}.(1 - v)^{1/2}.(1 - cv)^{1/2}.dv$$

Substituting $v = \mathrm{sn}^2(u, k)$ we obtain, after clearing,

$$z_1 = 2A.\int_0^K \mathrm{sn}^2u.\mathrm{cn}^2u.\mathrm{dn}^2u.du \quad (4)$$

Hence (see Appendix)

$$z_1 = (2A/15c^2).[2E(1 - cc') - c'K(2 - c)] \quad (5)$$

Taking the second set of values of the elements, we obtain $z = F(+7/2, +3/2, +3, c)$ which gives the integral

$$z_2 = 2A.\int_0^K \mathrm{sd}^2u.\mathrm{cd}^2u.\mathrm{nd}^2u.du \quad (6)$$

on substituting in equation (3) and putting $v = \mathrm{sn}^2(u, k)$. When integrated, this gives (see Appendix).

$$z_2 = (2A/15c^2c'^2).[2E(1 - cc') - c'K(2 - c)] \quad (7)$$

Thus z_1 and z_2 obtained this way are not independent solutions and are not satisfactory. On inspecting the other twenty-three solutions of equation (1) given by Forsyth[4] we see that $z = F(\alpha, \beta, \alpha + \beta - \gamma + 1, c')$ is given, and using the second set of values of the elements we obtain $z = F(+7/2, +3/2, +3, c')$ which gives rise to the integral

$$z_3 = 2A.\int_0^{K'} \mathrm{sd}^2u.\mathrm{cd}^2u.\mathrm{nd}^2u.du \quad (8)$$

when we put $v = \mathrm{sn}^2(u, k')$ and k' is the complementary modulus. This integral gives (see Appendix)

$$z_3 = (2A/15c^2c'^2).[2E'(1 - cc') - cK'(2 - c')] \quad (9)$$

This is a new solution and hence we can use z_2 and z_3 in the bilinear transformation equation, obtaining

$$z = \frac{A_1[2E(1 - cc') - c'K(2 - c)] + A_2[2E'(1 - cc') - cK'(2 - c')]}{B_1[2E(1 - cc') - c'K(2 - c)] + B_2[2E'(1 - cc') - cK'(2 - c')]} \quad (10)$$

In this, A_1, A_2, B_1 and B_2 are four constants which can be found from the corner conditions as follows:

(*a*) At the corner E, $z = 0$ when $c = 1$, and $E = 1$, $K = \infty$, $E' = \frac{1}{2}\pi$ and $K' = \frac{1}{2}\pi$. Hence $A_1 = 0$.

(*b*) At the corner F, $z = i.\infty$ when $c = 0$ and $c' = 1$ so that $K = \frac{1}{2}\pi$, $K' = \infty$, $E = \frac{1}{2}\pi$ and $E' = 1$. Hence $B_2 = 0$ since $\lim_{c \to 0} c(2 - c')K' = 0$.

Writing D for A_2/B_1, equation (10) becomes

$$z = D.[c(2 - c)K' - 2(1 - cc')E']/[c'(2 - c)K - 2(1 - cc')E] \quad (11)$$

At the corner B, $z = -h$ when $c = \infty$, where $\frac{1}{2}h$ is the radius of the semi-circular end BDE. To use this condition we must rewrite equation (11) in the form valid for when c is greater than unity. To do this we make the following substitutions: For c write $1/c$, for c' write $-c/c'$, for K write $k(K + iK')$, for K' write $k.K'$, for E write $(E - c'K - iE' + icK')/k$ and for E' write E'/k. We then obtain, after clearing,

$$z = \frac{-D.[(2c + c').cK' - 2(c^2 + c').E']}{cc'(2c - 1).(K + iK') + 2(c^2 + c')(E - c'K - iE' + icK')} \quad (12)$$

Putting $c = 0$ and $K = \frac{1}{2}\pi$, $K' = \infty$, $E = \frac{1}{2}\pi$, $E' = 1$, $c' = 1$ in equation (12) gives

$$z = -D/i = -h \quad \text{so that} \quad D = ih$$

Therefore the completed geometrical transformation is

$$z = ih.[c(2 - c')K' - 2(1 - cc')E']/[c'(2 - c)K - 2(1 - cc')E] \quad (13)$$

For $+1 \geqslant c \geqslant 0$ the integrals K, K', E and E' take on real values so that z will be a real multiple of ih for these values of c. These values are correct and correspond to the line EF of Fig. 1. For values of c greater than unity, the values of z will be complex. If $c = 2$ so that $c' = -1$, $E' = 1{\cdot}9103$, $K' = 1{\cdot}31101$, $E = 0{\cdot}5995(1 - i)$ and $K = 1{\cdot}31101(1 + i)$ giving $z = \frac{1}{2}h(i - 1)$ which are the co-ordinates of the point D, the mid-point of the circular arc BDE.

The values of z obtained from equation (13) are given below, and when plotted they lead to the correct curve of Fig. 1 showing that the geometrical transformation is correct. The real c-axis is shown in Fig. 2, and it can be seen from the

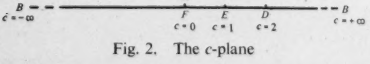


Fig. 2. The c-plane

distribution of the c-values along it that the area external to the edge $ABDEF$ of the conductor in the z-plane does correspond to the upper half of the c-plane.

THE ELECTRICAL TRANSFORMATION

It is now necessary to set up a one-to-one correspondence between the z-plane and the W-plane, where $W = U + i.V$ and is a complex potential. This is done by a transformation between the W-plane and the c-plane so that the real axis of the c-plane is given a potential of $+V_0$ e.s.u. in the electrical case (say). This, in effect, raises the potential of the edge of the conductor in the z-plane to $+V_0$ e.s.u. It is also necessary for the point at infinity in the z-plane to have a potential of $-\infty$ e.s.u., and this is achieved, as explained before, by the electrical transformation

$$dW/dc = -A/c^2 \quad (14)$$

THE ELECTRICAL INTENSITY R

The electrical (or magnetic) intensity (field strength) R is given by

$$R = |dW/dz| = |(dW/dc).(dc/dz)| \quad (15)$$

Now $dW/dc = -A/c^2$ so we require dz/dc from equation (13). In equation (13) write $N = c(2 - c')K' - 2(1 - cc')E'$ and $D = c'(2 - c)K - 2(1 - cc')E$ so that $z/ih = N/D$. Hence

$$(1/ih).(dz/dc) = (D.dN/dc - N.dD/dc)/D^2 \quad (16)$$

To evaluate this differential we use the following relations given by Cayley,[5] noting that $dk/dc = 1/2k$,

$$dE/dc = (E - K)/2c$$
$$dE'/dc = (K' - E')/2c'$$
$$dK/dc = (E - c'K)/2cc'$$
$$dK'/dc = (cK' - E')/2cc'$$

Hence $dN/dc = (5/2)(cK' + E' - 2cE')$

and $dD/dc = (5/2)(E - 2cE - c'K)$

Therefore $D.dN/dc - N.dD/dc = -15cc'.\pi/4$

using Legendre's relationship.

Hence, from equation (16)

$$(1/ih).(dz/dc) = -(15ih\pi cc')/[4c'(2 - c)K - 2(1 - cc')E]^2$$

This gives,

$$R = (4A/15\pi h)|[c'(2 - c)K - 2(1 - cc')E]^2/c'c^3| \quad (17)$$

RESULTS

Equation (17) gives the value of R at any point in the z-plane. Computation is laborious and only the variation of R along the sides FE and EDB and along the axis of symmetry DG (see Fig. 3) will be found.

(1) *The field strength R along the edge EF.* Along EF the modulus c is real and less than unity except at the point E where $c = 1$. The values of z/h are those given in Table 1.

Table 1. *Values of R along the edge EF* (see also Fig. 3)

c	=	1·000	0·8885	0·7500	0·5868	0·5000
z/ih	=	0·000	0·0095	0·1333	0·542	1·000
$R(15\pi h/4A)$	=	∞	36·75	15·84	8·44	6·48
c	=	0·410	0·3289	0·2500	0·0676	0·000
z/ih	=	1·975	3·68	7·51	82·6	∞
$R(15\pi h/4A)$	=	4·56	3·52	2·53	1·674	0·000

These results give the portion of the curve labelled EF in Fig. 3, where R in multiples of $(4A/15\pi h)$ is plotted against

the distance z measured from the point E (the origin) in multiples of ih. It can be seen that R varies from infinity at the corner E to zero at the point at infinity in the z-plane. The variation is smooth and what would be expected.

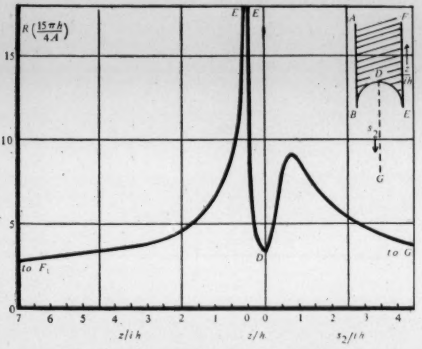


Fig. 3. The variation of R along $FEDG$

(2) *The variation of R along the curved end ED.* Along ED the modulus is real and varies from unity at the point E to $c = k^2 = 2$ at the point D. From symmetry the variation of R from D to B will be the same as that from E to D and hence will not be calculated. The corresponding values of z/h are those given in Table 2.

Table 2. *Values of R along the curved end ED*
(see also Fig. 3)

c	=	1·000	1·137	1·334
z/h	=	−0·000 +0·000i	−0·00043 +0·0175i	−0·0192 +0·118i
$R(15\pi h/4A)$	=	∞	26·46	8·724
c	=	1·703	2·000	
z/h	=	−0·209 +0·396i	−0·500 +0·500i	
$R(15\pi h/4A)$	=	4·02	3·22	

These results give the portion of the curve labelled ED in Fig. 3. It can be seen that R varies smoothly from an infinite value at the sharp corner E to a minimum value at the point D.

(3) *The variation of R along the centre line DG.* The centre line is shown dotted in the inset of Fig. 3 and is the external axis of symmetry of the pole-piece. Inspection of Fig. 2 shows that the curve in the c-plane corresponding to DG is a semi-circle joining the points $c = 2$ and $c = 0$ so that the values of c along DG will be complex. To simplify the calculations a second c-plane, which we shall call the c_2-plane, is introduced so that the line DG becomes the negative imaginary axis of this plane, and the area external to the pole-piece in the z-plane corresponds to the lower half of this plane. The transformation $c = 2/(1 + c_2)$ introduces the required symmetry, the point D now corresponding to the origin of the c_2-plane. Any point on the line DG has a corresponding c_2-value of $c_2 = -i.\alpha$ where α is a real variable. Now if $\alpha = \tan(\frac{1}{2}x)$ we find that $c = 1 + \exp(ix)$, where x is in right angles. Tables for the elliptic integrals K and K' for complex moduli of the form $c = 1 + \exp(ix)$ have been published by Cambi,[6] and the values of E and E' calculated from these tables are given by Langton.[7] The calculations involved in finding R and the associated values of z are very laborious and were made with the aid of a slide-rule (20 in.). For this reason the values of R and z are given to only three figures, and the last figure may be inaccurate. Since a knowledge of the variation of R along the line DG is the data most likely to be useful from the practical point of view, these values have been calculated for most of the values of x given by Cambi. The values of z should be of the form $z = a.h - 0{\cdot}500h$, where a is a positive or negative imaginary number, since these values of z are measured from the point D along the line DG. The full values of z are given in the table below so that an idea can be obtained of the accuracy of the calculations by observing how closely the second term agrees with the value $-0{\cdot}5000h$. It will be seen that most of the values are within 2% of the correct value, the worst values being when x is large, as then the calculations involve the difference of two nearly equal terms. The values of R are of a rather better order of accuracy. Values of s_2 are also given in Table 3, these being the distances in terms of h measured along DG from the point D where $s_2 = 0$, and have been obtained by ignoring the last term in the values of z, and subtracting $0{\cdot}5000ih$ from the first factor. Values of R in terms of the constant $(4A/15\pi h)$ are plotted against s_2/ih in Fig. 3. Here it can be seen that R rises to a maximum as we move along the line DG from the point D. This maximum occurs, as would be expected, nearly opposite the corners B and E, after this R steadily decreases, at first rapidly, and then more slowly, to become zero at the point at infinity.

Table 3. *Variation of R along DG*
(see also Fig. 3)

z/h	=	0·500i −0·500	0·377i −0·508	0·266i −0·498	
c	=	2·000	1·9877 +0·1564i	1·9510 +0·3090i	
s_2/ih	=	0·000	−0·124	−0·234	
$R(15\pi h/4A)$	=	3·22	4·06	5·08	
z/h	=	0·186i −0·509	0·109i −0·503	0·037i −0·499	−0·034i −0·502
c	=	1·890 +0·4580i	1·808 +0·5878i	1·707 +0·707i	1·588 +0·810i
s_2/ih	=	−0·314	−0·391	−0·463	−0·534
$R(15\pi h/4A)$	=	5·96	6·86	7·60	8·19
z/h	=	−0·106i −0·502	−0·183i −0·495	−0·262i −0·500	−0·358i −0·501
c	=	1·454 +0·890i	1·309 +0·951i	1·1564 +0·9877i	1·000 +1·000i
s_2/ih	=	−0·606	−0·683	−0·762	−0·858
$R(15\pi h/4A)$	=	8·67	8·98	9·21	9·18
z/h	=	−0·467i −0·500	−0·623i −0·498	−0·791i −0·499	−1·000i −0·491
c	=	0·8434 +0·9877i	0·6910 +0·9511i	0·5458 +0·890i	0·4122 +0·809i
s_2/ih	=	−0·967	−1·123	−1·291	−1·504
$R(15\pi h/4A)$	=	8·93	8·53	7·89	7·25
z/h	=	−1·465i −0·499	−2·152i −0·496	−3·505i −0·643	−7·61i −0·692
c	=	0·293 +0·707i	0·192 +0·5878i	0·11 +0·4580i	0·0488 +0·3090i
s_2/ih	=	−1·965	−2·65	−4·00	−8·11
$R(15\pi h/4A)$	=	6·16	5·09	4·00	2·66

REFERENCES

(1) Davy, N., and Langton, N. H. *Quart. J. Mech. Appl. Math.*, **6**, p. 115 (1953).
(2) Davy, N., and Langton, N. H. *Brit. J. Appl. Phys.*, **3**, pp. 156–158 (1952).
(3) Langton, N., and Davy, N. *Brit. J. Appl. Phys.*, **4**, pp. 134–137 (1953).
(4) Forsyth, A. R. *A Treatise on Differential Equations*, 3rd ed., pp. 210–212 (London: Macmillan and Co. Ltd., 1903).
(5) Cayley, A. *An Elementary Treatise on Elliptic Functions*, 2nd ed., p. 48 (London: G. Bell and Sons Ltd., 1895).
(6) Cambi, E. *J. Math. Phys.*, **26**, pp. 234–235 (1948).
(7) Langton, N. H., Ph.D. Thesis, University of Nottingham (1952).

APPENDIX

(1) *Evaluation of the integral* $z = 2A\int_0^K \mathrm{sn}^2 u.\mathrm{cn}^2 u.\mathrm{dn}^2 u.du.$

Writing the integral in terms of sn u, we obtain,

$$z/2A = \int_0^K \mathrm{sn}^2 u.du - (1+c)\int_0^K \mathrm{sn}^4 u.du - c.\int_0^K \mathrm{sn}^6 u.du$$

From Glaisher,*

$$(n+1)\int (k\,\mathrm{sn}\,u)^{n+2}du - n(1+c)\int (k\,\mathrm{sn}\,u)^n du + (n-1)c\int (k\,\mathrm{sn}\,u)^{n-2}du - k^n.\mathrm{sn}^{n-1}u.\mathrm{cn}\,u.\mathrm{dn}\,u = 0$$

If $n = 4$,

$$5\int_0^K c^3\,\mathrm{sn}^6 u.du - 4(1+c)\int_0^K c^2\,\mathrm{sn}^4 u.du + 3c\int_0^K \mathrm{sn}^2 u.du - \left[c^2.\mathrm{sn}^3 u.\mathrm{cn}\,u.\mathrm{dn}\,u\right]_0^K = 0$$

Now $\left[c^2.\mathrm{sn}^3 u.\mathrm{cn}\,u.\mathrm{dn}\,u\right]_0^K = 0$

Hence,

$$\int_0^K \mathrm{sn}^6 u.du = (1/5c^3).\left[4(1+c)\int_0^K c^2.\mathrm{sn}^4 u.du - 3c^2\int_0^K \mathrm{sn}^2 u.du\right]$$

If $n = 2$,

$$3\int_0^K c^2.\mathrm{sn}^4 u.du - 2(1+c)\int_0^K c.\mathrm{sn}^2 u.du + c.\int_0^K du = 0$$

Hence,

$$\int_0^K \mathrm{sn}^4 u.du = (1/3c^2).\left[2(1+c)\int_0^K c.\mathrm{sn}^2 u.du - c.K\right]$$

Now $$\int_0^K \mathrm{sn}^2 u.du = (K-E)/c$$

Therefore $$\int_0^K \mathrm{sn}^4 u.du = (1/3c^2).[2(1+c).(K-E) - c.K]$$

Hence we obtain, after rearranging,

$$\int_0^K \mathrm{sn}^6 u.du = (1/15c^3).[K(8+4c^2+3c) + E(-7c-8-8c^2)]$$

Substituting for the three integrals in the expression for z gives, after rearranging,

$$z/2A = (1/15c^2).[2E(1-cc') - c'.K(2-c)]$$

(2) *Evaluation of the integral* $z = 2A.\int_0^K \mathrm{sd}^2 u.\mathrm{cd}^2 u.\mathrm{nd}^2 u.du$

This integral comes from the equation

$$z = \mathrm{A}.\int_0^1 v^{1/2}.(1-v)^{1/2}.(1-c.v)^{-7/2}.dv$$

by putting $v = \mathrm{sn}^2(u, k)$.

Writing the integral in terms of nd u gives

$$z/2A = (1/c^2).\left[-c'\int_0^K \mathrm{nd}^6 u.du + (1+c')\int_0^K \mathrm{nd}^4 u.du - \int_0^K \mathrm{nd}^2 u.du\right]$$

Glaisher* gives,

$$5c'\int \mathrm{nd}^4 u.du - 4(1+c')\int \mathrm{nd}^4 u.du + 3\int \mathrm{nd}^2 u.du + c.\mathrm{nd}^3 u.\mathrm{cd}u.\mathrm{sd}u = 0$$

and

$$3c'\int \mathrm{nd}^4 u.du - 2(1+c')\int \mathrm{nd}^2 u.du + \int du + c.\mathrm{nd}\,u.\mathrm{cd}\,u.\mathrm{sd}\,u = 0$$

By a similar procedure to that outlined above, and noting that the elliptic functions are all to modulus k, and that $\left[\mathrm{nd}^3 u.\mathrm{cd}\,u.\mathrm{sd}\,u\right]_0^K = \left[\mathrm{nd}\,u.\mathrm{cd}\,u.\mathrm{sd}\,u\right]_0^K = 0$, we obtain

$$\int_0^K \mathrm{nd}^6 u.du = (1/15c'^3).[E(8+7c'+8c'^2) - 4c'K(1+c')]$$

and $$\int_0^K \mathrm{nd}^4 u.du = (1/3c'^2).[2E(1+c') - c'K]$$

using the fact that $$\int_0^K \mathrm{nd}^2(u,k).du = E/c'$$

Therefore, substituting in the expression for $z/2A$, we obtain, after rearranging,

$$z/2A = (1/15c^2c'^2).[2E(1-cc') - c'K(2-c)]$$

(3) *Evaluation of the integral* $z = 2A.\int_0^{K'} \mathrm{sd}^2 u.\mathrm{cd}^2 u.\mathrm{nd}^2 u.du.$

This integral comes from the equation $z = F(+7/2, +3/2, +3, c')$ by substituting $v = \mathrm{sn}^2(u, k')$. The elliptic functions

* Glaisher. *Messenger of Mathematics*, Vol. XI–XII, p. 126 (London: Macmillan and Co. Ltd., 1882).

are now to modulus k' and we are integrating between the limits of K' and 0, so that the result will be the same as that for the integral just evaluated except that in the new solution dashed symbols will be undashed and undashed ones will be dashed. Thus,

$$\int_0^{K'} \mathrm{nd}^2(u, k')du = E/c'$$

$$\int_0^{K'} \mathrm{nd}^6(u, k').du = (1/15c^3).[E'(8 + 7c + 8c^2) - 4cK'(1 + c)]$$

and

$$\int_0^{K'} \mathrm{nd}^4(u, k').du = (1/3c^2).[2E'(1 + c) - c.K']$$

giving, finally,

$$z/2A = (1/15c^2c'^2).[2E'(1 - cc') - c.K(2 - c')]$$

The mechanical impedance of the human mastoid process

By R. S. Dadson, M.A., D. W. Robinson, B.Sc., A.C.G.I., A.M.I.E.E., and R. G. P. Greig, B.Sc., National Physical Laboratory, Teddington, Middlesex

[*Paper received* 21 *June*, 1954]

At the instance of the Electro-Acoustics Committee of the Medical Research Council, the National Physical Laboratory has undertaken a comprehensive series of measurements of the driving-point impedance of the mastoid process of the head on twenty normal subjects for a variety of operating conditions. The main object of the work was to develop objective procedures for assessing the response characteristics of bone-conduction receivers, by providing data for the design of an "artificial mastoid" having a driving-point impedance simulating the average human mastoid. The design of such a device is in hand and will be reported in due course. A frequency range of 125–6200 c/s was covered which is rather wider than that currently utilized in hearing aids or audiometers operating by bone conduction. In the frequency band concerned the impedance of the average mastoid was found to be expressible simply as the impedance of the tissue layer backed by a rigid mass representing the inertia of the whole head. Equivalent electrical networks derived from the test results illustrate the dynamical properties of the system and their dependence on operating conditions such as the static contact force and the driven area. A short review of the small amount of comparable data reported by other authors is included.

1. INTRODUCTION

Hearing aids employing bone-conduction receivers have been in use for many years, and there is also a well-established technique of diagnostic audiometry in which the threshold of hearing by bone conduction has to be determined with a certain precision. Until comparatively recently, however, little has been known about the objective aspects of this subject and, in particular, reliable data for the rational design or calibration of bone-conduction receivers has been lacking.

The first important step in establishing a sound basis for objective measurements on bone-conduction apparatus consists of the development of an "artificial mastoid" so designed as to simulate the average human mastoid process in the load impedance which it presents to bone-conduction receivers under the range of conditions encountered in practice, and to secure the agreement and general use of this instrument among the interested users. The function of such a device is analogous to that of the artificial ear, which, for various reasons, has received more attention. As will be seen, however, the problem of realizing the artificial mastoid is possibly not more difficult than in the case of the artificial ear; for example, the complications of wave transmission do not appear to enter the problem, at least up to 6000 c/s.

As part of a programme of work, under the auspices of the Electro-Acoustics Committee of the Medical Research Council, on the establishment of standards of measurement in connexion with audiometry and hearing aids,[1] the National Physical Laboratory has undertaken to make the necessary basic measurements; in this paper the part of the project dealing with the experimental determinations of the driving-point impedance of the average human mastoid is described. The development of an artificial mastoid is attained in principle once this data is established; its physical realization rests on purely objective measurement and dynamical principles. The actual design of an artificial mastoid based on the results described in this paper is in progress and will be reported in due course. Several empirical designs of artificial mastoid have been described at various times,[2–4] but none of these seem to have been based on experimental data on the impedance of the human mastoid process.

2. SCOPE OF INVESTIGATION

One factor which substantially affects the load impedance presented to a bone-conduction receiver is the static contact force between the head and the receiver. The area of contact and the profile of the surface also have some influence, and these various effects are considered later in this paper. It should perhaps be stressed that the present paper deals only with the mechanical impedance presented to the receiver and does not attempt to consider the subsequent mechanism by which the transmitted sound excites the sensation of hearing.

The present measurements have covered the frequency range 125–6200 c/s, although current practice in bone-conduction audiometry and hearing-aid design is usually confined to a rather narrower band of frequencies. As will be seen, however, some interesting light is thrown on the dynamic mechanical properties of human tissue by the analysis of the impedance results at the extremities of the range, which would otherwise be liable to be overlooked. Bárány,[5] for instance, treats the tissue as exhibiting normal elasticity by defining for it a single compliance value, whereas the actual mechanism is considerably more complicated. This point will be treated more fully in Section 5.

The impedance measurements were made on the right

mastoid process of twenty normal subjects, including both men and women. The subjects were found to be representative of a much larger group in such respects as the dimensions of the head, the apparent thickness of the skin, and the exterior profile of the mastoid process. Two different contact styli were used (see Figs. 4 and 5), one having an elongated shape of $3 \cdot 5 \times 1 \cdot 9$ cm chosen as typifying the shape of bone conductors in current use on hearing aids. The other, representing the type of stylus used on some audiometers, was circular in form with a diameter of $1 \cdot 4$ cm. Except for a slight radius at the edges, both the styli were flat, the respective surface areas being $5 \cdot 88$ and $1 \cdot 54$ cm^2.

The styli were applied to the mastoid processes of the subjects with contact forces covering the range 200 to 500 g.wt for the "hearing-aid" type, and 500 to 1250 g.wt for the "audiometer" type, the principal sets of results being obtained at 400 and 1000 g.wt respectively. In the case of the "hearing-aid" stylus the longer axis was aligned with the ridge of the mastoid process. As the forehead is occasionally used as the driving point in clinical studies, measurements were also made in this position, using the "audiometer" stylus with a contact force of 1000 g.wt. The mechanical impedance to vibration normal to the surface was determined under this range of conditions for each of the subjects at a large number of frequencies by the methods described below.

3. METHOD OF MEASUREMENT

The method of mechanical impedance measurement most suitable for the present purpose seemed to be the determination of the ratio of the unknown impedance to that of a standard impedance. The most convenient form of the latter is a rigid vibrating mass, which, in the absence of appreciable air damping, acoustic radiation or internal wave motion, exhibits a calculable pure mechanical reactance. The driving apparatus consists of specially designed electro-mechanical transducers, arranged so that the comparison is made entirely in terms of electrical quantities. The transducer is equipped with means for observing the vibrational velocity at the driving point, that is, at the junction of the transducer and the unknown impedance. It is, of course, essential to ensure that the mode of vibration of the mechanical parts of the transducer results in motion of the driving point in the desired direction, in the present case parallel to the axis of the stylus. This point presents some difficulty in the design of the transducer when, as in the present case, it has to withstand a relatively large static force whenever the unknown impedance is applied. This results in some conflict between the factors of experimental sensitivity and mechanical rigidity. Sensitivity requires that the mechanical source impedance of the transducer should be at most of the order of the unknown impedance, in the present case of the order 10^4 mechanical ohms. On the other hand considerable mechanical rigidity is required in order to avoid alterations in the mode of vibration due to distortion of the system under the static stresses. The type and range of the transducer are largely dictated by these two factors.

The general arrangement adopted uses a form of massive resonant transducer, and is shown schematically in Fig. 1. In order to achieve adequate sensitivity, measurements are made at or near the frequencies of mechanical resonance of the transducer, that is to say, in the notation of Fig. 1, when the impedance of m_0 with S_0 is low enough in relation to the unknown impedance Z. The actual transducers employed are shown in Fig. 2. The larger unit shown at the left was designed for operation at the lower end of the frequency range. It consists of a large energized magnet and driving coil attached to which is a connecting rod which drives the vibrating bar, supported on a heavy casting, seen at the front. A number of such bars were employed having resonant

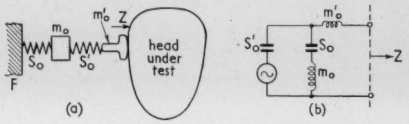


Fig. 1. (*a*) Simplified schematic diagram of mechanical system

F, reference frame (at rest); m_0, mass of transducer body; m_0', mass of vibrating stylus; S_0', stiffness of stylus suspension; S_0, stiffness of transducer mounting; Z, mechanical impedance under test.

Fig. 1. (*b*) Electrical network analogue of system

The electro-magnetically generated vibromotive force acts between m_0 and the vibrating system $m_0'S_0'$. m_0' and S_0' are approximately in resonance. S_0 has a large compliance.

frequencies from about 100 to 1000 c/s. The stylus is attached to an extension of the connecting rod, and interposed between it and the vibrating bar is the detachable standard mass consisting of a brass disk. Two forms of velocity indicator were used, one of which consisted of a coil secured to the connecting rod and centred in the gap of a freely suspended ring magnet. The frequency range of the larger instrument was limited at the upper end by dimensional considerations and also by diminished sensitivity resulting from rising internal mechanical impedance. For the measurements at frequencies greater than about 1000 c/s the second transducer seen at the right of Fig. 2 was employed. The velocity

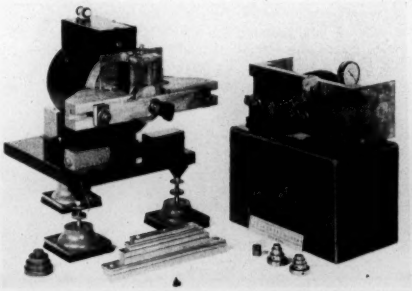


Fig. 2. Electromechanical transducers

indicator in this model was in the form of a wafer of barium titanate attached to a steel disk fixed as close as possible to the driving stylus. Moreover, in order to increase the sensitivity obtainable, a balanced vibrating bar system was employed on the same general principle as the ordinary tuning fork. By this means a substantially lower damping was achieved than with the single clamped-clamped bar of the low-frequency transducer. The driving coil and the stylus are connected to opposite members of the balanced bar system. Similar arrangements were made to vary the resonant

frequency and to attach the standard masses. The order of magnitude of the mechanical impedance of the two transducers at a few typical resonance frequencies is shown in Table 1.

Table 1. *Source impedance of transducers at resonance*

Low-frequency transducer		*High-frequency transducer*	
Frequency c/s	*Mechanical ohms*	*Frequency c/s*	*Mechanical ohms*
125	$0{\cdot}12 \times 10^4$	1400	$1{\cdot}36 \times 10^4$
492	1·86	4100	3·15
960	5·44	4800	6·28

The transducers were mounted on resilient supports of large compliance, the deflexion of which under the static contact force was used as a convenient means of standardizing the latter. To this end a calibrated dial gauge was mounted together with a mirror system (not shown in the photograph) in such a way that the subject under test was easily able to maintain the desired deflexion.

Two rather different electrical circuits have been used in the present investigation, one of which has a slight theoretical advantage while the other is more rapid and convenient. These are shown in Fig. 3. Let

Z_0 = mechanical source impedance of transducer,
Z = unknown mechanical impedance,
m = standard mass,
A = force factor of driving magnet system,
$\omega = 2\pi f$ = angular frequency,
$\dot{x}$ = vibration velocity,
E = output electromotive force of velocity indicator,
I = driving-coil current.

The first method [Fig. 3(*a*)] consists of observing $\dot{x}$ and I under the following four conditions:

(*a*) no load impedance $A.I_1 = Z_0.\dot{x}_1$
(*b*) standard mass attached $A.I_2 = (Z_0 + i\omega m).\dot{x}_2$
(*c*) unknown impedance applied $A'.I_3 = (Z_0' + Z).\dot{x}_3$
(*d*) unknown impedance applied in presence of standard mass $A'.I_4 = (Z_0' + Z + i\omega m).\dot{x}_4$

In conditions (*c*) and (*d*) the static load is present. As this may influence the internal geometry a little, the corresponding equations show A and Z_0 modified to A' and Z_0'. It can be seen that the variation, if any, in A can be eliminated from the equations. Unfortunately this is not the case with Z_0, since any variation in this quantity will appear as an inseparable component of Z, the measured impedance. However, it was possible to show indirectly that the variation in Z_0, as well as in A, was negligible. This was done by measuring the mechanical impedances of standard masses with the axis of the transducer alternately horizontal and vertical. In this way the force of gravity was made to simulate the application of a static load without introducing any change in the load impedances which are due simply to the inertia of the masses.

It is experimentally convenient to equate I_1, I_2, I_3 and I_4 by monitoring the potential difference across a resistor in series with the driving coil, and adjusting the output of the oscillator if necessary. Since, for given frequency, E is in each case proportional to $\dot{x}$ the equations reduce to

$$Z = -i\omega m.\frac{(E_3/E_4) + (E_1/E_2)}{(1 - E_3/E_4)(1 - E_1/E_2)}$$

The determination of the voltage ratios E_3/E_4 and E_1/E_2 was made by means of a measuring amplifier and a double-beam oscilloscope, giving accuracies of about $\pm 0{\cdot}1$ db in magnitude, and about 1° in phase angles, respectively.

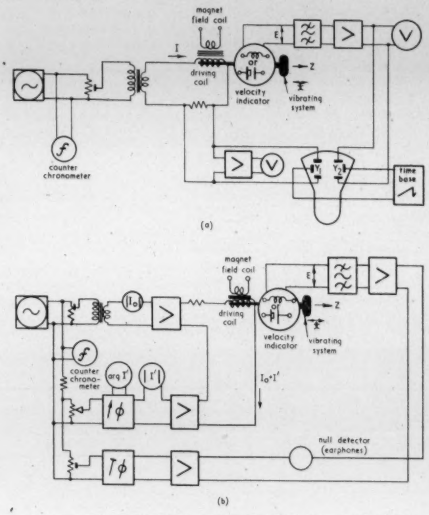


Fig. 3. Block schematic diagrams of electrical circuit arrangements
(*a*) Two-ratio method.
(*b*) One-ratio method.

The alternative method, shown in Fig. 3(*b*), is an adaptation of a method used by Wigan[6] for acoustic impedance measurement. The assumption $A = A'$ which is implied has already been seen to involve negligible error. The source of driving-coil current is replaced by two non-interacting current sources, one of which (I_0) is maintained constant, while the other (I') is adjustable in phase and magnitude. The vibration velocity is monitored by the velocity indicator and is initially set for the condition of no load impedance. When an external load impedance is applied, the current I' is introduced to restore the previous velocity indication. Three conditions of test are thus required:

(*a*) no load impedance $A.I_0 = Z_0.\dot{x}_1$
(*b*) standard mass attached $A(I_0 + I_1') = (Z_0 + i\omega m)\dot{x}_2$
(*c*) unknown impedance applied $A(I_0 + I_2') = (Z_0 + Z)\dot{x}_3$

By equating the three velocities the value of Z is found simply as $Z = i\omega m.I_2'/I_1'$. Thus, only a single observation of current ratio is required. This was determined in magnitude and phase by a measuring amplifier and phase shifter respectively. For the latter purpose, the resolver magslip device described by Patchett[7] was used. With this arrangement the required phase difference ($\arg I_2' - \arg I_1'$) could be read directly off the magslip protractor scale with an accuracy estimated at better than 1° at frequencies from 125–7000 c/s, provided the

resistance-capacity values in the stator circuit were accurately matched for each frequency in use. This estimate was based on comparisons with the double-beam oscilloscope method and an independent phase meter. A null method was used to maintain $\dot{x}$ constant. The output of the velocity indicator was balanced against a signal taken from the master oscillator through a preset alternating current potentiometer consisting of a resistive attenuator and a second resolver magslip phase controller.

Routine checks of the linearity and the intercomparison of the measured mechanical impedances of a range of standard masses were carried out with both methods for each frequency tested. These tests indicated an overall accuracy of about $\pm 2\%$ in magnitude of the measured impedance and $\pm 1°$ in phase.

4. RESULTS OF MEASUREMENTS

The results of the measurements are shown in Figs. 4 to 8. Expressing the mechanical impedance Z as $R + iX$ in mechanical ohms, the results given are in each case the mean of R and the mean of X for the group of subjects under the conditions indicated. The principal sets of experimental results are those shown in Figs. 4, 5 and 6 in which the force

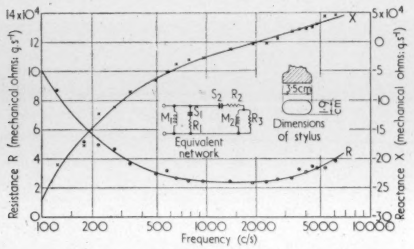


Fig. 4. Mechanical impedance $(R + iX)$ of mastoid process to "hearing-aid type" stylus; force of application 400 g.wt

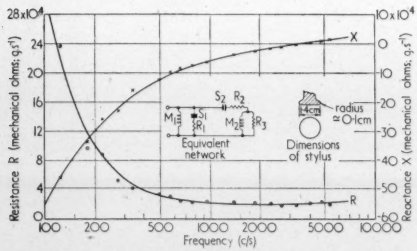


Fig. 5. Mechanical impedance $(R + iX)$ of mastoid process to "audiometer type" stylus; force of application 1000 g.wt

of application was 400 and 1000 g.wt for the "hearing-aid" and "audiometer" styli respectively. Figs. 7 and 8 show the dependence of the form of the impedance functions on the force of application for each of the styli. All the curves shown in Figs. 4 to 8 represent computed impedance functions fitting the experimental data by a least-squares criterion. Details of the method of computation are given in the Appendix.

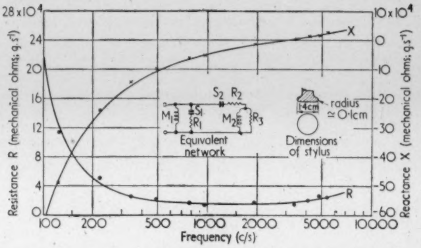


Fig. 6. Mechanical impedance $(R + iX)$ of forehead to "audiometer type" stylus; force of application 1000 g.wt

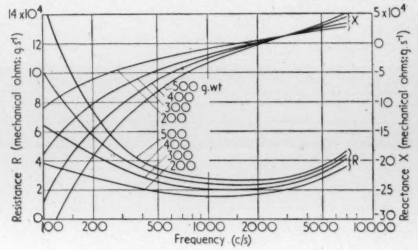


Fig. 7. Mechanical impedance $(R + iX)$ of mastoid process to "hearing-aid type" stylus; effect of varying force of application

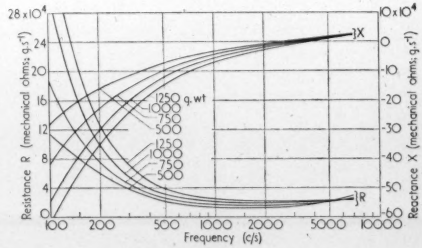


Fig. 8. Mechanical impedance $(R + iX)$ of mastoid process to "audiometer type" stylus; effect of varying force of application

The mean impedance for a group of subjects under any one condition of measurement is seen to be a smooth function of the frequency, and changing the form of the stylus or the force of application modifies the function in detail rather than in general form. Results for individual subjects follow in general the same type of curve as the means. The resistive component R shows a broad minimum in the middle of the

frequency range, whereas the reactive component rises steadily from negative to positive values, passing through zero in the same region. It is notable that the results for the mechanical impedance of the forehead (Fig. 6) are very similar to those for the mastoid process under similar test conditions (Fig. 5). As will be seen later, the mean results are also substantially the same for male and female subjects.

The measurements were made using r.m.s. velocities of the order 10^{-2} cm/s, but the impedance values were found to be independent of the amplitude over a range of about ± 20 dB relative to this level.

Broadly speaking the dispersion of the results at any one frequency was proportional to the modulus of the impedance at that frequency, and is therefore best expressed in terms of the ratio of the r.m.s. standard deviation for R and X to the modulus of Z. This ratio, averaged over all the tests, was 0·42, and it increased from 0·31 at the lowest frequency to 0·45 at the highest. The dispersion of the results with the "audiometer" stylus is slightly less than with the other type. Bearing in mind that the results relate to a group of twenty subjects, the statistical data may be summed up briefly by stating that the fiducial limits for Z at the probability level 0·98 are approximately $\pm 20\%$. This figure contains components due to the inherent differences between the mastoids of different subjects as well as the experimental errors. It was possible to analyse the variance into these two components in fifteen cases, for which the results shown are actually the mean of two separate determinations by the same group. The results were found to be fairly stable and indicated that about 40% of the total variance was due to individual differences and 60% to "error." The latter component includes not only the observational and instrumental errors, but also such factors as the variability of position of the stylus on the same mastoid in repeated tests.

Some tests on the variation of impedance with position of application showed that there was no systematic variation over a distance of some 3 cm along the ridge of the mastoid process, but at distances of the order 1 cm off the ridge the impedance values tended to be lower by some 40%.

5. EQUIVALENT NETWORKS AND ANALYSIS OF RESULTS

The treatment of the impedance data by analytical methods leads naturally to their representation by equivalent electrical networks, by means of which the operation of various factors is clarified. The equivalent network representation implies that the impedance functions shall either be, or be approximated by, a rational algebraic function of the complex frequency parameter. Accordingly a method was developed of determining the form of such functions of pre-assigned degree to approximate sufficiently closely to the experimental results. The method used was an extension of the interpolating procedure given by Carlin[8] and is described in full in the Appendix. The computational work was undertaken by the Mathematics Division of the Laboratory, employing the Automatic Computing Engine (ACE). A separate evaluation was carried out for the following sets of experimental results:

"Hearing-aid" stylus; 20 subjects; 200, 300, 400 and 500 g.wt;

"Audiometer" stylus; 20 subjects; 500, 750, 1000 and 1250 g.wt;

"Hearing-aid" stylus; 400 g.wt; male and female data treated separately;

"Audiometer" stylus; 1000 g.wt; male and female data treated separately;

"Audiometer" stylus applied to forehead; 20 subjects; 1000 g.wt.

The resulting functions are illustrated by the curves in Figs. 4–8.

The functions representing the experimental data may be synthesized into a variety of equivalent networks by known techniques.[9–12] A form which facilitates identification between the network elements and the physical features is illustrated in Fig. 9(*a*) and is obtained by a partial-fraction expansion of the admittance function. The parameters so determined, and their dependence on force of application, are shown in Fig. 10, and some supplementary values in Table 2.

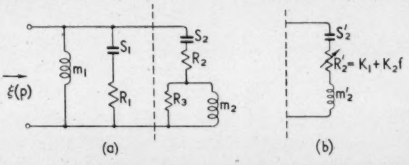


Fig. 9. Electrical networks equivalent to mechanical impedance of mastoid or forehead

(*a*) Normal form.
(*b*) Alternative form of third branch.

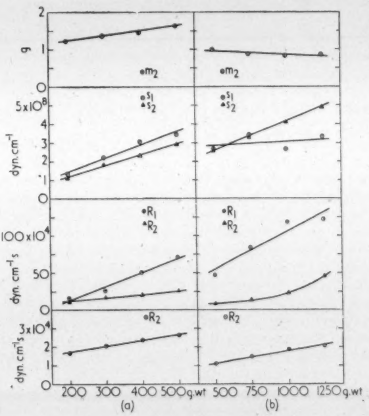


Fig. 10. Variation of network parameters with contact force for two types of stylus

(*a*) Hearing-aid stylus.
(*b*) Audiometer stylus.

The mass term m_1 represents the mass of the whole head and assumes a value of about 3·6 kg. This may be compared with the "dynamic mass" of 3·0 to 3·5 kg quoted by Békésy.[13] The stiffness, damping and inertia of the tissue layer are represented by the third branch. These items, together with m_1, constitute the simplest representation in terms of lumped mechanical constants which could be

expected to give a reasonable approximation to the actual data.

The existence of the centre branch, the presence of which is indispensable for accurate representation over the whole frequency range studied, clearly shows that the above simple version neglects important properties of the mechanical system. It seems that S_1 and R_1 must be taken to represent the phenomenon known as "retarded elasticity," which is known to occur in other resilient materials such as many high polymers. The time constant or retardation time, which is equal to R_1/S_1, has a value of a few milliseconds. If an approximating function of higher degree were found, the resulting network would probably be similar but would contain additional (R, S) branches representing other retardation times. In the limit this would lead to a distribution of infinitesimal branches of this type corresponding to a continuous distribution of retardation times as is known to occur in certain high polymers and rubber-like materials.[14-17] For instance, the retardation time for polyisobutylene at 25° C given by Ferry, Grandine and Fitzgerald[17] has a distribution with a fairly sharp maximum at 5 ms. This latter refinement is, however, unnecessary for the main purposes of the present work, and would require an impracticably high precision in the experimental results to justify it.

Just as the centre branch of the network approximates a continuous distribution by means of two lumped constants, so the branch m_1 is also a simplified representation. The head is not in an exact sense an inertia terminal, since this ignores its attachment to the body and the possibility of some relative internal motion. However, the resolution of this branch into a sub-network containing a plurality of elements would certainly require accurate experimental data at frequencies much lower than 125 c/s, which would not be relevant to the normal use of bone-conduction instruments.

The configuration of the third branch (S_2, R_2, M_2, R_3), representing normal elasticity and inertia of the tissue, also requires some comment. The form shown in Fig. 9(*a*) is that which followed naturally from the curve-fitting process, on the assumption of constant network elements. It can be shown, however, that this branch may be closely approximated by a series-resonant network of only three elements provided frequency-dependent damping is allowed. [See Fig. 9(*b*).] If one assumes for simplicity a resistive element of the form $R = k_1 + k_2 . f$, where k_1 and k_2 are constants, the appropriate constants may be determined by a further application of the process described in the Appendix. The data for the "hearing-aid" stylus with 400 g.wt contact force lead, for example, to the following parameters:

$$m_2 = 1{\cdot}41 \text{ g}$$
$$S_2 = 2{\cdot}25.10^8 \text{ dyn.cm}^{-1}$$
$$R = (2{\cdot}06.10^4 + 2{\cdot}46.f) \text{ dyn.cm}^{-1}\text{s}$$

in place of the values shown in the first row of Table 2.

Referring to Fig. 10, each of the parameters of the equivalent network, except m_1 which, in conformity with its physical significance, is nearly constant, is seen to increase with contact area. Each, with the further exception of m_2, also increases with contact force. These results may be interpreted as follows. The increase of the stiffness term with contact force results partly from increased stiffness of the tissue itself when more highly compressed, and partly from the increased contact area due to the squashing of the normally convex surface of the tissue by the flat surface of the stylus. The damping terms would be expected to vary with the stiffnesses. Any increase of m_2 with contact pressure due to increased contact area seems to be partly offset by the effect of lateral extrusion. In the case of the large stylus, m_2 shows a net increase with contact force, but with the "audiometer" type stylus the trend is reversed.

Table 2. *Parameters of equivalent networks*

Stylus	*Contact force g.wt*	*Group of subjects*	m_1 *kg*	m_2 *g*	$S_1 \times 10^{-8}$ *dyn.cm*$^{-1}$	$S_2 \times 10^{-8}$ *dyn.cm*$^{-1}$	$R_1 \times 10^{-4}$ *dyn.cm*$^{-1}$s	$R_2 \times 10^{-4}$ *dyn.cm*$^{-1}$s	$R_3 \times 10^{-4}$ *dyn.cm*$^{-1}$s
H.A.	400	20	3·6	1·47	3·06	2·33	50·3	2·39	20·3
H.A.	400	Males	3·6	1·71	2·63	2·19	52·6	2·34	24·0
H.A.	400	Females	3·6	1·32	3·55	2·77	38·9	2·38	16·0
Aud.	1000	20	3·6	0·81	2·64	4·08	118	1·86	22·9
Aud.	1000	Males	3·6	0·89	2·91	3·86	102	1·82	17·9
Aud.	1000	Females	3·6	0·75	2·43	4·50	139	1·79	16·0
Aud.	1000	20 (foreheads)	3·5	1·17	0·67	3·38	227	1·47	13·7

6. TABULATION OF NETWORK PARAMETERS AS A FUNCTION OF STYLUS AREA AND CONTACT FORCE

It may be convenient for comparative studies to express the parameters of the network of Fig. 9(*a*) as functions of the stylus area and contact force which are likely to be the most important variables.

It has been found that each of the network parameters can be expressed as a function linear in these two variables, with the exception, of course, of m_1 which is constant, and R_3 for which a quadratic variation in the force is required.

These functions have been evaluated by fitting the points shown in Fig. 10 by the least squares criterion. Expressing the parameter concerned in the form $Q = (a + bS) + (C + dS)F + (e + fS)F^2$, where S and F are respectively the area (cm^2) and the force of contact (kg.wt), Table 3 gives

Table 3. *Interpolating table for network elements*

	m_1 *g*	m_2 *g*	$S_1 \times 10^{-8}$ *dyn.cm*$^{-1}$	$S_2 \times 10^{-8}$ *dyn.cm*$^{-1}$	$R_1 \times 10^{-4}$ *dyn.cm*$^{-1}$s	$R_2 \times 10^{-4}$ *dyn.cm*$^{-1}$s	$R_3 \times 10^{-4}$ *dyn.cm*$^{-1}$s
a	3600	+1·09	+3·6	+1·39	+13·0	+0·26	+ 38
b	—	−0·0177	−0·61	−0·24	−6·4	+0·142	− 5·3
c	—	−0·73	−1·97	+2·1	+72	+0·67	−114
d	—	+0·34	+1·56	+0·64	+19·4	+0·42	+ 24
e	—	—	—	—	—	—	+ 89
f	—	—	—	—	—	—	− 11·8

the values of the six constants a to f for each of the seven network elements.

Allowing a reasonable degree of extrapolation, the data in Table 3 may be used to estimate the equivalent networks for values of the stylus area from about 1 to 7 cm^2 and of the contact force from 0·2 to 1·5 kg.wt, and hence to estimate the corresponding average impedance of human mastoids over the range of frequency 100–8000 c/s.

7. COMPARISON WITH OTHER DATA

A number of other workers have reported measurements of the mechanical impedance of the mastoid process under various conditions, but so far no complete account comparable in scope with the present work has been published. The amount of previous material available is thus small and, on account of the variety of conditions adopted, comparison with the present results is mostly rather indirect and needs careful interpretation.

Franke[18] has determined the mechanical impedance of five mastoids over the range of frequency 200 to 1800 c/s, and has obtained results which show the same general variation with frequency and which are consistent with our values, after allowing for differences of the stylus area and force of application, to within a factor of 2. Franke, however, incorrectly concludes that the reactance of the tissue is purely elastic, due probably to the somewhat narrow frequency range covered by his measurements. He also considers it justifiable to regard the resistive component as constant, but this conclusion is not borne out by the present results.

Some comparatively sparse data have been given by Bárány[5] and Békésy.[13] Values given by Bárány for the resistance and equivalent stiffness at 435 and 1024 c/s under several test conditions are, on the average, consistent with our results to within about 20%. As previously remarked, Békésy suggests a value for the dynamic mass of the head in close agreement with our value for m_1 and, working at 30 c/s, has also obtained measurements which, with a certain amount of extrapolation, lead to values of the stiffness of the tissue in quite good agreement with our data. Comparisons between our data and those of the other authors mentioned were facilitated by the use of Table 3, from which it is possible to estimate the results we should have obtained under the test conditions of the other workers.

Preliminary reports have also appeared indicating that work has recently been carried out at the Post Office Research Station[19] and the National Bureau of Standards[20]; details of this work will presumably become available in due course.

8. CONCLUSIONS

(*a*) With the main object of providing a basis for the objective assessment of the performance of bone-conduction devices, the mechanical impedance of the average human mastoid has been determined for a group of twenty normal subjects. The measurements were made with two styli of different surface areas and a range of contact force, and covered the frequency range 125 to 6200 c/s. Analysis indicates that the data may safely be extrapolated to at least the range 100 to 8000 c/s.

(*b*) The measured impedances are found to be smooth functions of the frequency, whether for individuals or for the mean, and are similar in general form over the range of test conditions covered.

(*c*) All the results indicate that the impedance may be regarded as that of the tissue layer backed by a rigid mass having a value corresponding to the total mass of the head. Analysis by means of equivalent networks shows that the behaviour of the tissue is consistent with that of a material exhibiting both normal and time-dependent elasticity, the latter showing a retardation time of a few milliseconds.

(*d*) On the basis of the above analysis it has been possible to fit smooth impedance curves to the experimental data and to draw up formulae for the network parameters as functions of the stylus area and force of application. These enable the impedance values to be easily computed over a wide range of conditions.

(*e*) The results appear to form a suitable basis for the design of an "artificial mastoid" for the calibration of bone-conduction receivers of the types used for hearing aids or audiometers. The design of such a device is proceeding and will be described in due course in another paper.

ACKNOWLEDGEMENTS

The authors acknowledge the patient co-operation of the subjects who took part in the long series of tests and the help of Mr. E. G. Butcher and Mr. J. H. King in the circuit arrangements and measurements. Valuable assistance was given by the Mathematics Division of the Laboratory in the extensive computations involved in the network analysis. Thanks are also due to the Royal Aircraft Establishment, Farnborough, for the loan of a phase meter. The work described in this paper was carried out as part of the research programme of the National Physical Laboratory and the paper is published by permission of the Director of the Laboratory.

REFERENCES

(1) DADSON, R. S. *Proc. Int. Electroacoust. Cong. Netherlands* 1953, Section IV, p. 151.
(2) HAWLEY, M. S. *Bell Lab. Record*, **18,** p. 73 (1939).
(3) GREIBACH, E. H. *Trans Amer. Inst. Elect. Engrs*, **65,** p. 184 (1946).
(4) CARLISLE, R. W., and PEARSON, H. A. *J. Acoust. Soc. Amer.*, **23,** p. 300 (1951).
(5) BÁRÁNY, E. *Acta Otolaryngologica*, Suppl. XXVI–XXVII (1938).
(6) WIGAN, E. R. Private communication (1948).
(7) PATCHETT, G. N. *Electronic Engng*, **23,** p. 310 (1951).
(8) CARLIN, H. J. *Proc. Inst. Radio Engrs*, **37,** p. 744 (1949).
(9) BRUNE, O. *J. Math. Phys.*, **10,** p. 191 (1930).
(10) BOTT, R., and DUFFIN, R. J. *J. Appl. Phys.*, **20,** p. 816 (1949).
(11) GUILLEMIN, E. A. *Advances in Electronics III*, p. 271 (New York: Academic Press, Inc., 1951).
(12) FRY, T. C. *Bull. Amer. Math. Soc.*, **35,** p. 463 (1929).
(13) BÉKÉSY, G. v. *J. Acoust. Soc. Amer.*, **20,** p. 749 (1948).
(14) KUHN, W., and KÜNZLE, O. *Helv. Chim. Acta*, **30,** p. 839 (1947).
(15) ALFREY, T. *Mechanical Behaviour of High Polymers* (New York: Interscience Publishers, Inc., 1948).
(16) TRELOAR, L. R. G. *The Physics of Rubber Elasticity* (London: Oxford University Press, 1949).
(17) FERRY, J. D., GRANDINE, L. D., and FITZGERALD, E. R. *J. Appl. Phys.*, **24,** p. 911 (1953).
(18) FRANKE, E. K. *J. Acoust. Soc. Amer.*, **24,** p. 410 (1952).
(19) MORTON, J. Y. Discussion to Ref. (1).
(20) CORLISS, EDITH L. R., and KOIDAN, W. *J. Acoust. Soc. Amer.*, **26,** p. 141 (1954). (Abstract only.)

APPENDIX

Procedure of network analysis

The process of fitting the impedance data by smooth functions of the frequency, leading to the synthesis of

equivalent electrical networks, is based on the method outlined by Carlin[8] in another connexion. The applications of this method to the fitting of experimental data do not appear to be widely known and an outline of the steps in the procedure adopted for the results in this paper may be of interest.

The experimental impedance data are regarded as sample values of an impedance function $Z(p)$ say, where $p = \sigma + i\omega$ is the complex angular frequency parameter. In the present application $Z(p)$ relates to a system containing living tissue and may therefore possess internal sources of energy. It is, however, considered extremely unlikely that such energy is concerned in the response of the system to external vibration. $Z(p)$ has therefore been assumed to have the properties of the impedance function of a passive system, and this assumption is justified by the interpretation we have been able to give of the elements of the networks so derived.

$Z(p)$ is, therefore, presumed to be a positive real function of undetermined degree n in p.

The actual vibrating system possesses potentially an almost infinite number of degrees of freedom, so that $Z(p)$ may be of very high degree or transcendental, and this means that it can only be determined from a finite set of experimental data provided some assumption about its mathematical form is imposed. The most convenient simplication is to assume that the number of degrees of freedom is effectively finite and even quite small. The validity of this restriction has to be tested by the goodness of fit to the experimental data. With n finite (equivalent to the restriction to "lumped constants"), the driving-point impedance $Z(p)$ will be approximated by a rational fractional function of the form

$$\xi(p) = \frac{a_1 p^n + a_2 p^{n-1} + \ldots + a_{n+1}}{(p - w_1)(p - w_2)\ldots(p - w_n)},$$

where n is for the moment arbitrary, and the constants a_r and the poles w_r are to be determined. The optimum value of n is found by equating it successively to 1, 2, 3 .... and carrying out the analytical procedure described below until the resulting function approximates the experimental data as closely as required. In the present case the proper value of n was quite well defined, and was equal to 4. With $n < 4$ it was either impossible to secure a fit to the data over the whole frequency range, or the function so found was not positive real, and with $n > 4$ some of the constants a_r and w_r were inordinately sensitive to the input data.

The essence of Carlin's method is a formal process for determining the constants in $\xi(p)$, which ensures that the function is potentially realizable, that is to say, its poles are either on the negative real axis of p or in conjugate pairs in the left half-plane and the constants in the numerator are real numbers. The process consists of evaluating the integral

$$\xi(p) = \frac{1}{2\pi i} \int_C \left[1 - \frac{(p-p_1)\ldots(p-p_{n+1})(w-w_1)\ldots(w-w_n)}{(w-p_1)\ldots(w-p_{n+1})(p-w_1)\ldots(p-w_n)}\right] \frac{Z(w)dw}{w-p}$$

which, by a theorem due to J. L. Walsh,[8] assumes the values $Z(p)$ at each of the poles of the integrand $p_1 \ldots p_{n+1}$, provided the contour C encloses all these poles. This expression imposes no restriction on the values $w_1 \ldots w_n$. By summing the residues of the integrand, $\xi(p)$ is obtained as a rational fractional function of p with $w_1 \ldots w_n$ occurring in each of the coefficients, and since these constants are confined to real or conjugate complex pairs, each numerator coefficient can be resolved into its real and imaginary parts. The condition of vanishing of the imaginary parts leads to explicit solutions for $w_1 \ldots w_n$, and hence to $\xi(p)$. This process of itself does not guarantee that the function is a realizable impedance function; a function so fitted to arbitrary data would not in general be realizable. In the present application, however, the data is not arbitrary; apart from experimental error, it consists of samples from an actual impedance function $Z(p)$. It is therefore to be expected that the function $\xi(p)$ will be positive real, provided n is sufficiently large.

The simplest application of the method, illustrated by examples in the reference quoted, occurs when the function $\xi(p)$ is made to assume numerical values for $(n + 1)$ values of the argument, one of which must be zero (or, by minor adjustments, infinity). Reference to the experimental values in the present case (for example, Figs. 4 to 6) shows that the function $Z(p)$ is varying rapidly at both ends of the frequency range, and consequently no *a priori* estimation of the value at zero or infinite frequency is permissible. A modification of the method which does not place this restriction on the values of p consists of introducing a dummy value for $Z(p)$ at one of the interpolating points. This can be illustrated for the simple case of $n = 1$, when $\xi(p)$ is of the form $(a_1 p + a_2)/(p - w_1)$. The presence of the three constants a_1, a_2, w_1 shows that three data may be fitted, for example, $p = p_1, \mathscr{R}Z(p) = x_1, \mathscr{I}Z(p) = y_1$; $p = p_2, \mathscr{R}Z(p) = x_2$, where x_1, y_1, x_2, p_1 and p_2 are known numerical values. The dummy value $y_2 = \mathscr{I}Z(p_2)$ is introduced and the integral for $\xi(p)$ evaluated. y_2 is eliminated in the manipulation. Another modification which may be used if numerical values of the derivative of $Z(p)$ are known for one or more values of p_2 is to replace the corresponding factors $(w - p_r)$ in the integrand by $(w - p_r)$,[2] since the evaluation of the residue at these poles involves the value of the first derivative of $Z(p)$ at $p = p_r$. Each of these variants was applied to the present data and indicated that a close approximation was obtained with $n = 4$. The constant term in the numerator was found to be rather unstable and so small that its vanishing had negligible influence on the function in the range of approximation, and it was concluded that the form of $\xi(p)$ fitting the data in the range 125 to 6200 c/s was

$$[a_1 p^4 + a_2 p^3 + a_3 p^2 + a_4 p]/(p + b_1)(p + b_2)(p + b_3)(p + b_4)$$

with all the constants positive. This stage of the approximation process was carried out using data from freehand smoothing of the experimental impedance values.

The form of the function having been found it was a logical step to make use of all the experimental data instead of samples selected at rather wide intervals of frequency. For this purpose, the Mathematics Division of the Laboratory devised an iterative computing programme for use with the Automatic Computing Engine (ACE), for fitting a function of the prescribed form to the complete set of experimental data for each condition of test. The criterion of least squares of the relative error was adopted, namely

$$\delta \sum_i k_i \left\{ \frac{[\mathscr{R}Z(p_i) - \mathscr{R}\xi(p_i)]^2}{|Z(p_i)|^2} + \frac{[\mathscr{I}Z(p_i) - \mathscr{I}\xi(p_i)]^2}{|Z(p_i)|^2} \right\} = 0$$

where k_i is a weight factor proportional to the number of observations per point.

The functions resulting were synthesized into a number of equivalent networks by various methods described elsewhere.[9-12] The partial fraction expansion of the admittance function $1/\xi(p)$ [shown in Fig. 9(*a*)] was found to be best adapted to physical interpretation.

The shape of the emission bands of luminescent solids

By C. C. VLAM, Ph.D., Kamerlingh Onnes Laboratory, University of Leiden

[*Paper received* 20 *May*, 1954]

Single emission bands of luminescent solids generally exhibit a bell-shaped form. Closer examination of emission curves of a simple nature demonstrates a very close Gaussian form, but only when energy is plotted as a function of frequency. More complex emission curves can be very well described as a sum of Gaussian distribution functions. The experimentally determined shape of the emission profile is compared with theoretical emission curves, calculated from the configuration co-ordinate model.

THE EXPERIMENTAL DETERMINATION OF THE BAND SHAPE

Visual inspection of emission curves of luminescent solids frequently suggests the existence of one or more bell-shaped bands. Borissow[1] analysed emission spectra of sulphide phosphors into symmetrical and bell-shaped component bands, each represented by a Gaussian error function in the wavelength scale. The same procedure was applied to sulphide phosphors by Steinbacher[2] and Kutzner.[3]

The emission spectrum was also plotted in other ways. Schellenberg,[4] Henderson[5] and Butler[6] considered the intensity distribution as a function of frequency, whereas Szigeti, Nagy and Makai[7] analysed curves where the number of photons was plotted *versus* frequency. In both cases single emission bands were considered to be Gaussian.

Only Steinbacher questioned the validity of the method, in particular whether a single emission band might be Gaussian when energy was plotted *versus* wavelength or frequency. The accuracy attained in the measurements did not allow him to choose between these possibilities. In principle, since a symmetrical analysing function has great advantages, a choice between the different methods depends on whether a single emission band is expected to be symmetrical when energy is plotted *versus* wavelength or frequency, or when the number of photons is plotted *versus* frequency.

In order to decide between the various procedures, the intensity distribution of a number of phosphors was measured with accuracy, initially for curves of a simple nature.[8] It appeared that the broad emission bands of the tungstates are only perfectly symmetrical when energy is plotted as a function of frequency. As an analysing function the Gaussian error curve was the most appropriate (Fig. 1). Variation of temperature causes the emission band to shift, the curve, however, retains the Gaussian shape almost exactly. Justification for the procedure is also shown by the analysis of the emission curve of a zinc beryllium silicate phosphor, where again the main band shows a Gaussian profile (Fig. 2). Hence, the energy distribution in the wave-number scale is given by $I(\nu) = A \exp[-B(\nu - \nu_0)^2]$, where $I(\nu)$ is the intensity at wave number ν. The parameters are directly connected with peak intensity (A), band width at half peak intensity ($H = 1{\cdot}665/\sqrt{B}$), peak wave number (ν_0) and total band intensity ($O = A\sqrt{\pi}/\sqrt{B}$).

Emission curves composed of two or more bands could be very well described as a sum of Gaussian curves. In the case of sulphide phosphors the bands may not overlap very much, but the component bands of the zinc silicates and halophosphates overlap to a much greater extent.

When emission bands overlap considerably the accuracy of the analysis is dependent on the accuracy attained in the measurements. Lonn[9] has shown that a curve composed of Gaussian curves can be resolved into such curves in only one definite way. Consequently, though the analysis of simple emission curves shows a Gaussian energy distribution, the validity of the procedure depends on the justness of the assumption that, in the more complex case, the component bands do not deviate noticeably from the Gaussian shape. Sometimes no direct evidence of this assumption can be

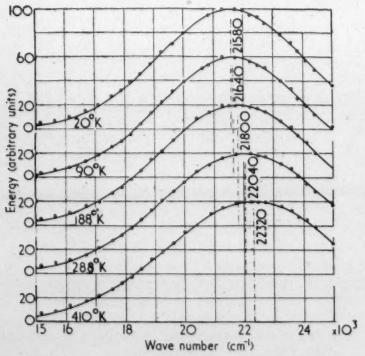


Fig. 1. Emission spectra of $CaWO_4 \cdot 1\%PbWO_4$ at different temperatures. Excitation by λ2537 Å

○ = measured points;
—— = Gaussian curves.

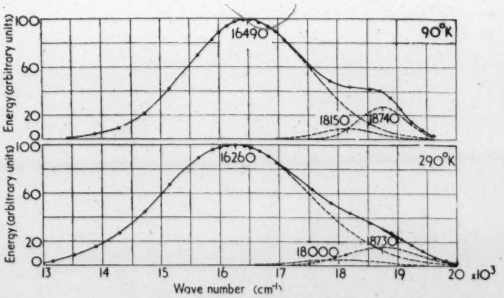


Fig. 2. Emission spectra of zinc-beryllium (84 : 12) silicate (4% Mn) at 90° K and 290° K. Excitation by λ2537 Å

○ = measured points; – – – = Gaussian curves;
—— = sum of the Gaussian curves.

given. In the case of the halophosphates, analysis demonstrates a strong overlap of the manganese and antimony bands, which is in full accordance with other data, i.e. the measured dependence of the decay on frequency.[10] Since the green emission of zinc silicate with low manganese content shows no dependence of decay constant on frequency, Klick and Schulman[11] have denied the existence of two separate bands in the emission spectrum. Actually, since analysis shows that the bands strongly overlap, the determination of a possible difference between the two decay constants is seriously hampered. Arguments in favour of the analysis into two bands are: (*a*) the total emission is accurately the sum of two Gaussian error curves; (*b*) the red emission band in manganese-activated zinc beryllium silicate is exactly Gaussian; (*c*) increase of the manganese content in zinc silicate clearly results in an increase of the low frequency emission, i.e. new bands appear or established bands increase in intensity; (*d*) the temperature dependence of the total emission for silicates with increasing manganese content indicates an increase of the number of emission bands;[12] (*e*) the temperature dependence of the intensity of each of the two sub-bands, obtained in analysis, is in fair agreement with the theoretical curves computed on the basis of the configuration co-ordinate model, assuming two excited states.[8] Physical reality can therefore be claimed for the existence of the subsidiary bands in manganese-activated zinc silicates.

Though mathematics may play an important role in the analysis, trial and error is in essence the basis of the analysing method used. Resolution of curves composed of two bands can generally be carried out unambiguously. In the case of three component bands, the resolution into individual bands is accurate when one of them is predominant (Fig. 2). At higher temperatures the analysis becomes more difficult on account of the greater width of the subsidiary bands. Thus the higher temperature curve of Fig. 2 cannot be resolved with great precision. Generally, in cases where the analysis is difficult, further information can be supplied by variation of the activator content. With the limitations mentioned, the analysis shows that the mathematical representation of emission bands by Gaussian error curves is a valuable tool. It allows us to give a simple analytical expression for the emission spectrum, and an adequate indication of the influence of factors such as activator content, temperature and method of excitation on the spectral distribution.

THE SHAPE OF A SINGLE EMISSION BAND ACCORDING TO THE THEORY

Two models have played the most important part in the explanation of luminescence phenomena, the zone model and the configuration co-ordinate model. In both cases the calculation of the intensity distribution within a single emission band requires precise information about the position of the initial levels and the transition probabilities involved. Also a reasonable weight function for the occupation of the different emission levels must be available. The zone model is less suitable for the calculation, since it cannot adequately account for the atomic vibrations and rearrangements taking place in the centre during absorption and emission. In the more "dynamical" configuration co-ordinate model the transition probabilities and the weight function for the excited states follow in a more fundamental way.

Assuming a one-co-ordinate description to be possible, in Fig. 3 the energy of the centre in the ground state as a function of this co-ordinate is denoted by $U^g(r)$, the analogous function for the excited state by $U^e(r)$. The minima are taken respectively as $r = 0$ and $r = r_0$. The intensity distribution of the emission is given quantum-mechanically by

$$I(\nu_{nm}) = CN_n\nu_{nm}^4\left[\int\psi_n^e(r - r_0)\psi_m^g(r)dr\right]^2 \qquad (1)$$

where C is a constant, N_n the number of centres in the nth vibrational state, $\nu_{nm} = (E_n^e - E_m^g)/hc$ = the wave number of the light emitted, E_n^e and E_m^g being the vibrational energies in the states with normalized eigenfunctions ψ_n^e and ψ_m^g respectively.

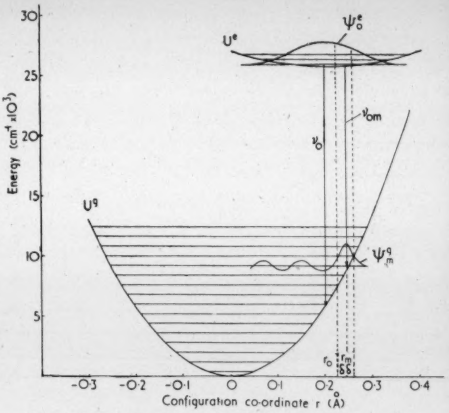


Fig. 3. Configuration co-ordinate diagram for $CaWO_4$.Pb: the energy of the WO_4^{2-} centre as a function of the W–O distance r. U^g and U^e are parabolic curves

Parameter values: $k_g = 5, 8 \cdot 10^5$ dyne/cm; $k_e = 1, 5 \cdot 10^5$ dyne/cm; $r_0 = 0,2$ Å and $U_0 = 25\,700$ cm^{-1}.

COMPARISON OF AN APPROXIMATE EXPRESSION FOR BAND SHAPE WITH EXPERIMENT

In the neighbourhood of the minima the potential energy curves may be considered to be parabolic, i.e. the oscillators are treated as harmonic. Consequently we have $U^g(r) = k_g r^2/2$ and $U^e(r - r_0) = U_0 + k_e(r - r_0)^2/2$ where k_g and k_e are force constants in the ground state and excited state of the centre respectively and U_0 is the energy difference between the minima of the potential energy curves. Equation (1) becomes

$$I(\nu_{nm}) = (C\nu_{nm}^4 N_n/\pi 2^{n+m}n!m!)\left\{\int H_n(y)H_m(z)\exp\left[-(y^2 + z^2)/2\right]dr\right\}^2 \qquad (2)$$

where $H_n(y)$ and $H_m(z)$ are Hermite polynomials and $y = 2\pi(r - r_0)(\mu\omega_e/h)^{\frac{1}{2}}$, $z = 2\pi r(\mu\omega_g/h)^{\frac{1}{2}}$ with μ the effective vibrating mass and ω_g and ω_e the classical frequency of the oscillator in the ground state and the excited state respectively.

Assuming temperature equilibrium is established before emission takes place, at low temperature all the excited centres are in the lowest vibrational state. If m is a high vibrational state, the product $\psi_0^e\psi_m^g$ in equation (1) is alternately positive and negative and r-values beyond the small interval $(r_m - \delta,\ r_m + \delta)$ do not contribute to the overlap integral. When the value of r_0 is relatively large we

can assume in a first approximation that $U^g(r)$ is a straight line in the interval where ψ_0^2 differs noticeably from zero. Putting $\nu_0 = [E_0^e - U^g(r_0)]/hc$ we finally obtain, neglecting the factor ν^4

$$I(\nu) = A \exp [-B(\nu - \nu_0)^2]$$

where $B = \gamma\sqrt{k_e/k_g^2}$ and $\gamma = 2\pi hc^2\sqrt{\mu}/r_0^2$

The width of the band follows from $H = 1{\cdot}665/\sqrt{B}$. Thus for low temperatures ($T \sim 0°$ K) the approximate treatment leads to Gaussian emission bands.

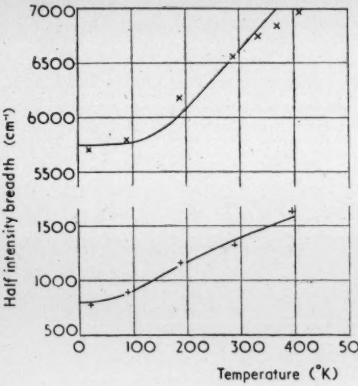


Fig. 4. Width of emission bands compared with theory [equation (4)]

Upper curve: theoretical curve with $H_0 = 5750$ cm^{-1} and $\omega_e/c = 398$ cm^{-1}; × = measured width of calcium tungstate band. *Lower curve:* theoretical curve with $H_0 = 810$ cm^{-1} and $\omega_e/c = 147$ cm^{-1}; + = measured width of main band of zinc silicate (0, 1% Mn).

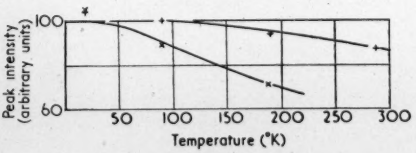


Fig. 5. Peak intensity of the emission bands compared with theory [equation (3)]

Upper curve: theoretical curve with $\omega_e/c = 398$ cm^{-1}; + = measured peak intensity of calcium tungstate. *Lower curve:* theoretical curve with $\omega_e/c = 147$ cm^{-1}; × = measured peak intensity of main band of zinc silicate (0, 1% Mn). Temperature ranges where no quenching occurs.

At higher temperatures the contribution of a few higher vibrational states has to be taken into account. The fraction of centres in the nth vibrational state is given by the Boltzmann factor

$$a_n = \exp [-E_n^e/kT] / \sum_{p=0}^{\infty} \exp [-E_p/kT]$$

The intensity is, assuming no loss by radiationless transitions, proportional to $\sum_{n=0}^{\infty} a_n \psi_n^2$. In the case of the harmonic oscillator the energy levels are equidistant and the eigenfunctions are normalized Hermite functions. The sum can be evaluated explicitly and we have

$$I = PQ_T \exp [-y^2 Q_T^2]$$

where $Q_T^2 = \tanh h\omega_e/2kT$ and P is a constant. Consequently for each temperature the profile of the emission bands remains Gaussian. The maximum intensity $A(T)$, for temperatures where no quenching occurs is given by

$$A(T) = A(0)[\tanh h\omega_e/2kT]^{\frac{1}{2}} \quad (3)$$

For each temperature the width $H(T)$ is given by

$$H(T) = H(0)[\coth h\omega_e/2kT]^{\frac{1}{2}} \quad (4)$$

In Figs 4 and 5 it is seen that the agreement of the theoretical curves with experiment is fair.

APPLICATION OF THE THEORY TO SPECIAL PHOSPHORS

In the previous section the shape of the emission spectrum was computed on the basis of the configuration co-ordinate model. The assumptions made, namely a relatively large value of r_0, and the neglect of a factor ν^4, lead to a Gaussian intensity distribution. More rigorously, the intensity distribution can be calculated from equation (2). In this way the emission spectrum of $CaWO_4$ was computed, assuming the WO_4^{--} ion to constitute the centre.[13] Only the radial vibrations of the four oxygen ions surrounding the tungsten ion were assumed to be of importance. Consequently μ is the mass of the oxygen ion; the force constants k_g and k_e for the vibrations in the ground state and the excited state were computed from the Raman oscillation spectrum and the variation of the band width with temperature respectively (Fig. 4). The value of r_0 was estimated as 0·20 Å; this value is slightly larger than the corresponding displacement computed for the free molecule OsO_4, which has the same electronic configuration. The parameter U_0, which only determines the peak frequency of the emission was accordingly chosen equivalent to 25 700 cm^{-1}. The result of the calculation is shown in Fig. 6.

It is noticeable that the fine structure predicted in equation (2) is obscured. This may be due to the broadening

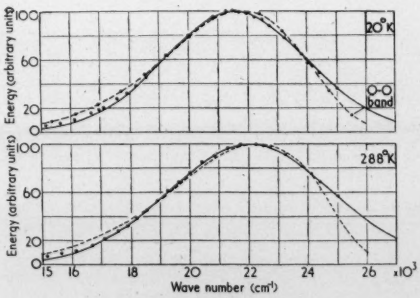


Fig. 6. Spectral energy distribution of calcium tungstate at 20° K and 288° K

—— ○ —— = measured points and best fitting Gaussian curve.
– – – ● – – – = calculated emission curves.

of the lines, caused by interaction with the lattice, Stark splitting of the levels, and lattice defects; moreover other modes of vibration may be of importance. For these reasons the observed fine structure may sometimes show an irregular pattern. The latter is the case with zinc silicate with low manganese content, where in the low temperature emission nineteen lines are observed which are not equidistant. The width of these lines never exceeds 10 Å. Therefore, Klick's calculation of this emission spectrum,[14] where the high frequency wing of the emission can only be justified by assuming a line width of approximately 200 Å, is very improbable. The discrepancy may arise from his doubtful assumption that the potential energy curves are parabolic up to the point where these curves intersect. Presumably the small width of the lines observed in the fine structure indicates that the O–O transition must have a small probability, as with thallium-activated potassium chloride, computed by Williams,[15] and with the tungstates. This being true, the r_0-value will be larger than follows from Klick's calculation, i.e. the preliminaries for Gaussian shaped sub-bands are better fulfilled. Possibly in the future more detailed investigations of the fine structure in emission spectra will clarify the problems of calculation of emission spectra, including the vibrational modes of the centre.

ACKNOWLEDGEMENT

The author wishes to acknowledge the help of Dr. S. T. Henderson, who corrected the English of the paper.

REFERENCES

(1) Borissow, P. *Ann. Phys.*, **42**, p. 1321 (1931).
(2) Steinbacher, P. Thesis, Heidelberg (1927).
(3) Kutzner, W. *Z. Phys.*, **106**, p. 551 (1937).
(4) Schellenberg, O. *Ann. Phys.*, **13**, p. 249 (1932).
(5) Henderson, S. T. *Proc. Roy. Soc. A*, **173**, p. 323 (1939).
(6) Butler, K. H. *J. Opt. Soc. Amer.*, **37**, p. 566 (1947). *Trans Electrochem. Soc.*, **91**, p. 13 (1947), **93**, p. 143 (1948). *J. Electrochem. Soc.*, **97**, p. 265 (1950).
(7) Szigeti, G., Nagy, E., and Makai, E. *J. Chem. Phys.*, **15**, p. 881 (1947).
(8) Brinkman H., and Vlam, C. C. *Physica*, **14**, p. 650 (1948).
Vlam, C. C. *Physica*, **15**, p. 609 (1949); *J. Opt. Soc. Amer.*, **41**, p. 558 (1951); Thesis, Groningen (1953).
(9) Lonn, E. *Z. Phys.*, **75**, p. 348 (1932).
(10) Studer, F. J., and Rosenbaum, A. *J. Opt. Soc. Amer.*, **39**, p. 685 (1949).
(11) Klick, C. C., and Schulman, J. H. *J. Opt. Soc. Amer.*, **40**, p. 509 (1950). *Phys. Rev.*, **85**, p. 154 (1952). *J. Phys. Chem.*, **57**, p. 776 (1953).
(12) Kröger, F. A. *Some aspects of the luminescence of solids*, p. 190 (Amsterdam: Elsevier, 1948).
(13) Hameka, H. F., and Vlam, C. C. *Physica*, **19**, p. 943 (1953).
(14) Klick, C. C., and Schulman, J. H. *J. Opt. Soc. Amer.*, **42**, p. 910 (1952).
(15) Williams, F. E. *J. Chem. Phys.*, **19**, p. 457 (1951).

High current spark channels

By J. E. Allen, B.Eng., Ph.D.,* and J. D. Craggs, M.Sc., Ph.D., F.Inst.P., Department of Electrical Engineering, The University of Liverpool

[*Paper received* 9 *April*, 1954]

Measurements have been made on high current (up to 265 kA) spark channels in air, argon and hydrogen. The voltage gradients existing in the channels were determined by oscillographic measurements and were found to increase with increasing current in the range studied. Power inputs of tens of megawatts per centimetre length of channel were derived from the measured voltage gradients. Rotating-mirror photographs of the discharges were taken, using a camera with a resolution better than 1/5 μs. The photographs obtained show a number of interesting features which are peculiar to these high current discharges, the chief one being the bright central core of the discharge, which was photographed in every case. Temperatures of about 10^5 °K were estimated from electron mobility considerations. The central core of the discharge appears to be due to the self-magnetic "pinch" effect, and calculations based on this inference suggest that high ion densities ($\sim 10^{19}$ per cm^3) exist in the centre of the core; the energy requirements for such conditions are compatible with the measured power inputs.

The term "spark channel" refers to the conducting path existing between the electrodes of a spark gap shortly after the initial breakdown processes have been completed. Meek and Craggs[1] have recently reviewed the experimental studies of spark channels carried out by various workers. Most of the previous investigations have referred to relatively low current discharges, but Norinder and Karsten[2] have studied high current sparks in air, in measurements limited to currents not greater than 102 kA; the measurements of Flowers[3] include currents up to 100 kA.

The high current generator[4] used in the present experiments[5] is being used to determine whether the self-magnetic "pinch" effect is appreciable in transient, high pressure, high current discharges[6] and to determine the temperatures attained.

Rotating-mirror photographs show that the discharges have a bright central "core"; they also show other features which support the view that this core is due to the "pinch" effect. Electron temperatures of about 10^5°K have been estimated from the observed voltage gradients and current densities, assuming that the electrons lose the momentum, which they gain from the electric field, mainly in electron-ion collisions. The positive ion temperature is expected to be nearly equal to the electron temperature in these high-pressure discharges, even in the short times involved (a few microseconds).

The theoretical contributions to the subject of magnetically constricted discharges[7–11] relate to steady-state, low-pressure arcs, hence they are certainly inapplicable to the transient, high-pressure sparks described in this paper. Experiments on the "pinch" effect in low-pressure transient discharges have recently been carried out by Ware[12] and Cousins and

* Now at the Atomic Energy Research Establishment, Harwell.

Ware.[13] Thonemann and Cowhig[10] have studied the effect in continuous discharges at low pressure.

APPARATUS

The high current generator which was used to produce the high current sparks has been described by Durnford and Reynolds;[4] it consists of a bank of 114 condensers (0·5 μF, 25 kV) which can be discharged through a circuit of low impedance. Similar generators have been described by Bellaschi,[14] McEachron and Thomason,[15] Norinder and Karsten,[2] Stekolnikov[16] and, very recently, Pfestorf.[17] A maximum current of 470 kA has been obtained using the present generator[4] and this has been equalled only by Stekolnikov.[16]

The inductance of the generator, which determines the discharge current (for a given amount of stored energy), depends to a large extent on the length of the spark gap assembly. The maximum current available with the assembly built for the present experiments is 265 kA. This assembly consists of a chamber containing the main spark gap and a superstructure containing two series spark gaps for triggering purposes. The spark chamber, which is 8 in. in diameter, has a glass window $5\frac{1}{2}$ in. in diameter and $\frac{1}{4}$ in. thick; the thick glass was used in order to withstand the impact of the shock wave propagated by the spark and the maximum pressure difference of 1 atmosphere, with a vacuum in the chamber.

The discharge current and voltage waveforms were recorded using a 20 kV continuously evacuated cathode-ray oscillograph. The current shunt is of the concentric cylinder type, with tapping points inside the inner cylinder.[4] It was connected to measure the current from a group of three condensers and the total current was obtained by multiplying by the appropriate factor. The resistance of the shunt is 0·06557 Ω and its performance is satisfactory at the frequency involved ($\sim$ 30 kc/s).

A rotating mirror was used to study the discharge, as in previous work by Beams[18] and others.[1] The spark channel was focused on to a slit mounted perpendicular to the channel, and the element of the channel image defined by the slit was scanned along a photographic plate in the direction of the image length. This is the usual streak camera technique.[19] The present camera has a polished steel mirror which can be rotated at speeds up to 450 rev/s, giving scanning velocities up to about 1·1 mm/μs. A slit width of nominal value 0·16 mm was used in the present work, giving a resolution better than $\frac{1}{2}$ μs. A small auxiliary mirror mounted on the same shaft, but in a light-tight compartment, reflects light from a small lamp on to the photocathode of a 931 A photomultiplier tube. The output pulses from this tube are used to trigger the generator at the correct instant and determination of their frequency provides a measure of the mirror rotation speed.

THE MEASUREMENT OF VOLTAGE GRADIENTS

The voltage gradients along high current spark channels in air, hydrogen and argon were determined by oscillographic measurement of the voltage drop across the gap for several gap lengths. The current waveform was oscillatory and reached its maximum value after 7·7 μs.

After many preliminary experiments the arrangement illustrated in Fig. 1 was used to obtain the final results. A small, single-stage impulse generator, charged to a voltage of mV_c (where V_c is the charging voltage and $m \simeq \frac{1}{2}$) was used to trigger the high current generator. The spark gaps of this small generator, G_a and G_b, were set so that

$$mV_c < V_{B.D.a} < mV_c + V$$

and

$$V < V_{B.D.b} < mV_c$$

where V is the magnitude of the negative triggering pulse. Thus, application of this pulse caused the successive breakdown of the gaps G_a and G_b; the small generator then triggered the cathode-ray oscillograph time base and caused a negative impulse wave of magnitude mV_c to be applied to the trigger electrode of the main generator (the magnitude is actually slightly less than mV_c, due to the stage efficiency of the small impulse generator). This wave reached its maximum value after about 0·6 μs, thus providing the necessary time for the oscillograph time sweep to start.

The main spark gaps were set, by trial and error, so that

$$V_c < V_{B.D.1} < V_c(1 + m)$$

$$mV_c < V_{B.D.2} < V_c$$

and

$$V_{B.D.3} < V_c$$

It was necessary to irradiate the spark gaps so that they would break down in succession.[1] A radioactive cobalt source, equivalent to about 10 mg of radium, was used for this purpose and was placed outside the spark chamber "in line" with the main gap (G_3).

One half of the spark gap voltage was applied to the deflexion plates of the oscillograph by means of a potential

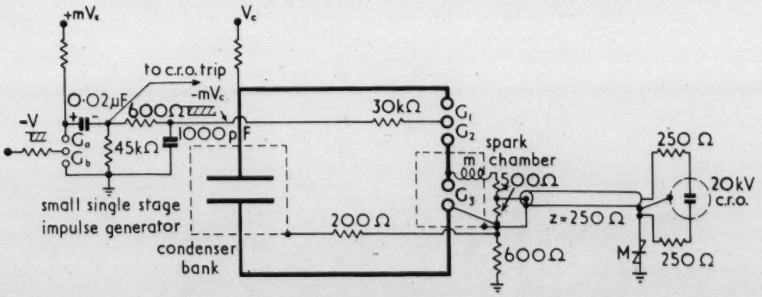


Fig. 1. Schematic arrangement of the circuit used to measure the voltage drop across high current spark channels. M is a metrosil disk and m is a mutual inductance coil

divider and a matched concentric cable. The oscillograph could withstand one half of the pre-breakdown voltage without damage, and the deflexions after breakdown were adequate for measurement. Thus the use of a non-linear divider as used by other workers [1] was avoided.

The earthing arrangements were critical: it was essential to bring all earth connexions to one point and to earth that point *via* as short a lead as possible. Oscillations due to the "capacitance to case" of the condensers and the inductance of the earth lead were damped out by the 200 Ω resistor shown in Fig. 1. Smaller oscillations, due to "capacitance to earth," were damped out by the 600 Ω resistor in the main earth lead.

The voltage oscillograms were distorted by an induced voltage resulting from the larger values of di/dt and for this reason the voltage drop was measured at the instant when this induced e.m.f. is zero, i.e. at the current maximum when $di/dt = 0$. To obtain oscillograms suitable for accurate measurement (i.e. of small slope dv/dt when $t = 7{\cdot}7\ \mu s$) the induced voltage in the measuring circuit was reduced by subtracting the voltage which was developed across a mutual inductance coil (m in Fig. 1). The voltage across this coil due to its self-inductance was quite negligible. The inductance of the discharge varies during the first few microseconds[2] because the diameter varies during this period. As a result it was not possible to cancel completely the induced voltage and obtain an oscillogram showing the discharge voltage as a function of time.

Experiments were performed, using different gap lengths and currents, in (1) air at atmospheric pressure, (2) hydrogen at 1 lb/in.^{2} gauge, and (3) argon at 1 lb/in.^{2} gauge. The electrodes were tungsten cylinders 15 mm in diameter. Measurements[5] on sparks in air indicate that the discharge voltage (at $t = 7{\cdot}7\ \mu s$) is independent of the electrode material, except when the latter is of relatively low boiling point.

ROTATING-MIRROR PHOTOGRAPHY

Rotating-mirror photographs were taken of 188 kA sparks in air, hydrogen and argon (gap length = 5·55 mm) and 265 kA sparks in air (gap length = 11·1 mm). Correct synchronization, so that the image of the discharge appeared on the film or plate, was obtained by adjustment of the optical system mentioned above.

Photographs were taken on Super XX film and Ilford H.P.3 plates and some of the films were calibrated using a step wedge.

EXPERIMENTAL RESULTS

(*a*) *Voltage gradients.* The voltage gradients, which refer to $t = 7{\cdot}7\ \mu s$ (when $di/dt = 0$), were obtained by differentiating voltage drop-gap length curves of the type illustrated by Fig. 2. The gradients are highest near the electrodes and gap lengths greater than 5 mm are required for the gradient in the middle of the gap to become independent of the gap length; the values of these gradients are given in Table 1, together with values of the power input/cm length. The measurements show that the voltage gradients are high and increase with increasing current in the range studied (188–265 kA). See also Fig. 3. The experimental error in these voltage gradient measurements is difficult to assess.

(*b*) *Rotating-mirror photographs.* Typical photographs are shown in Figs. 4 and 5; photographs obtained at different scanning speeds are shown, but the period of oscillation is the same in all cases (30·8 μs). The photographs are complex and the principal features are listed below, and shown in Fig. 6.

(1) The velocity of the initial expansion is high, always higher than the velocity of sound in the ambient gas.

(2) A luminous wavefront is always associated with this rapid expansion.

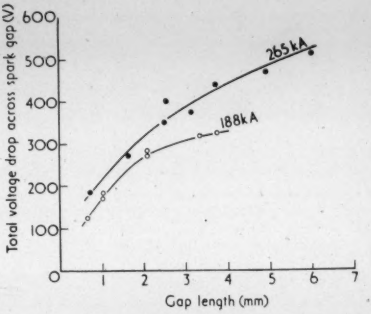


Fig. 2. Voltage drop-gap length curves for sparks between tungsten electrodes in air

Table 1. *Voltage gradients and power inputs*

	Current (kA)	*Voltage gradient (V/cm)*	*Power input/cm length (MW/cm)*
Air	188	180	34
	265	360	95
Hydrogen	188	350	66
	265	1000	265
Argon	188	80	15

Norinder and Karsten[2] found a power input of 23 MW/cm, at a peak current of 102 kA, when the voltage gradient was 230 V/cm. The quarter period was 12 μs.

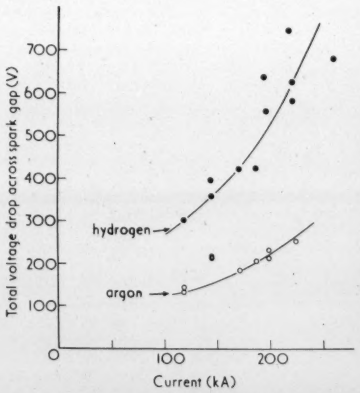


Fig. 3. Voltage-current curves for hydrogen and argon sparks between tungsten electrodes 4 mm apart

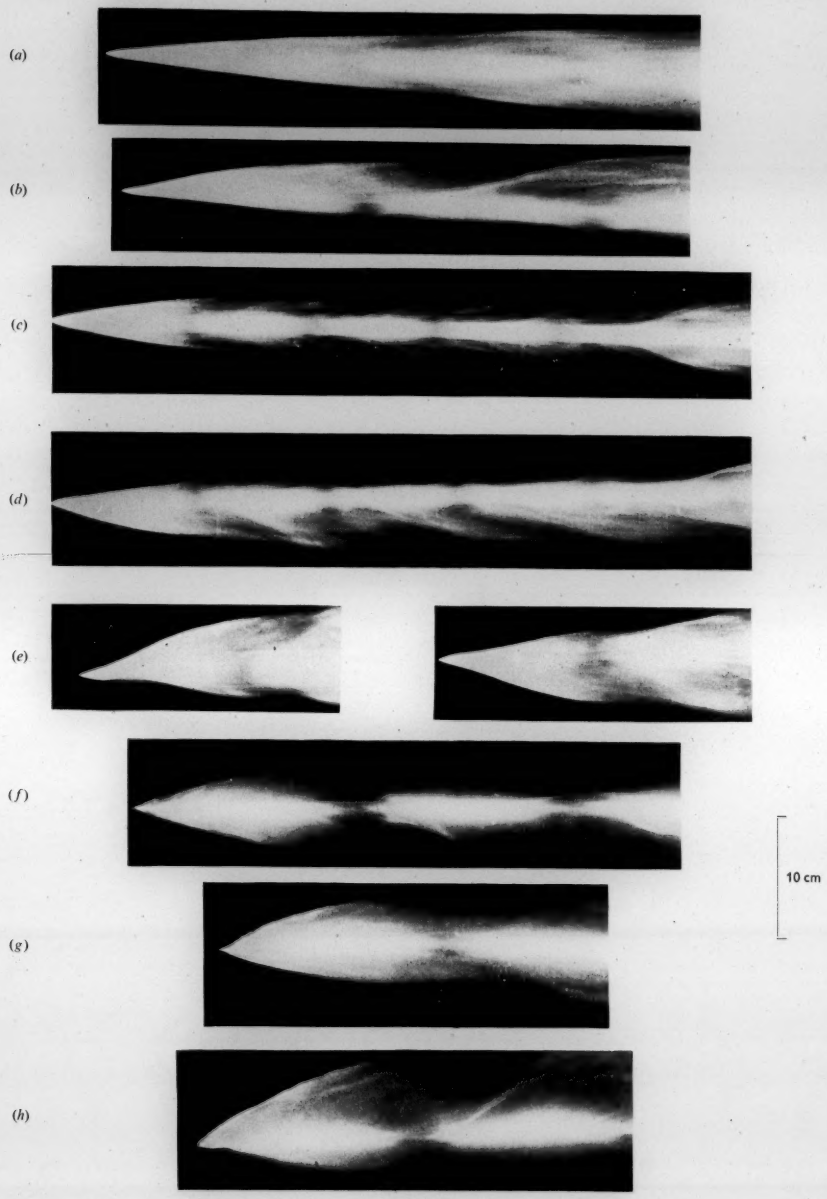


Fig. 4 (for legend see over)

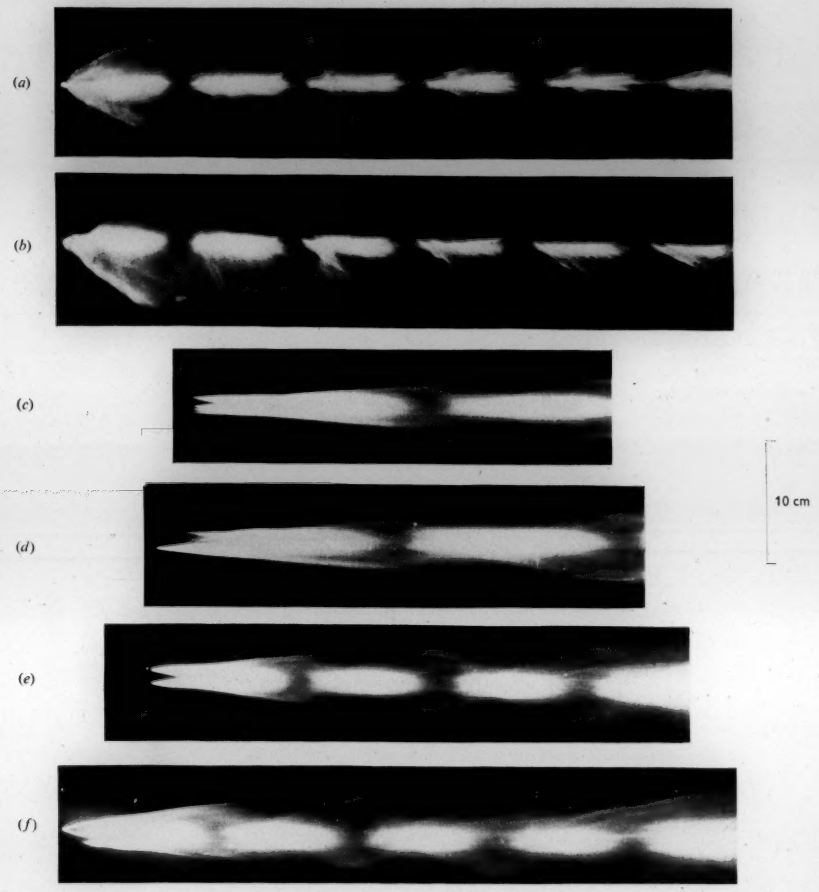


Fig. 5. Typical rotating-mirror photographs

(*a*) Symmetrical 188 kA spark in hydrogen; mirror speed 238 rev/s.
(*b*) Asymmetrical 188 kA spark in hydrogen; mirror speed 238 rev/s.
(*c*) 188 kA spark in argon; mirror speed 454 rev/s. Five channels can be seen in the initial stages.
(*d*) 188 kA spark in argon; mirror speed 454 rev/s.
(*e*) 188 kA spark in argon; mirror speed 302 rev/s.
(*f*) 188 kA spark in argon; mirror speed 302 rev/s.

Fig. 4 (see p. 449). Typical rotating-mirror photographs

(*a*) Symmetrical 188 kA spark in air; mirror speed 396 rev/s.
(*b*) Asymmetrical 188 kA spark in air; mirror speed 396 rev/s.
(*c*) Symmetrical 188 kA spark in air; mirror speed 302 rev/s.
(*d*) Asymmetrical 188 kA spark in air; mirror speed 302 rev/s.
(*e*) 265 kA sparks in air; mirror speed 285 rev/s.
(*f*) Symmetrical 188 kA spark in hydrogen; mirror speed 447 rev/s.
(*g*) 188 kA spark in hydrogen; mirror speed 447 rev/s. Two cores appear to exist during part of the first quarter cycle.
(*h*) Asymmetrical 188 kA spark in hydrogen; mirror speed 447 rev/s.

(3) Most of the records obtained were asymmetrical (examples of both symmetrical and asymmetrical photographs are given in Figs. 4 and 5).

(4) A bright central "core" exists in all cases.

(5) Sometimes, near the end of a current half cycle, a part of the "core" moves out from the centre but returns as the current builds up again in the opposite direction [see Figs. 4(*f*, *g*) and 5(*a*)].

(6) Inward-moving "waves" are very often seen towards the end of a half-cycle. For example, Fig. 4(*a*) shows such a wave towards the end of the second half-cycle and Fig. 4(*b*) shows one towards the end of the first half-cycle. Fig. 5(*f*) clearly shows how the second inward wave is due to a reflexion of an outward moving wave at the channel boundary.

(7) High velocity "jets" can be seen issuing from the central core [see, for example, Fig. 5(*b*)].

A schematic streak record illustrating the above features is shown in Fig. 6. Table 2 contains some numerical data, derived from the rotating mirror photographs.

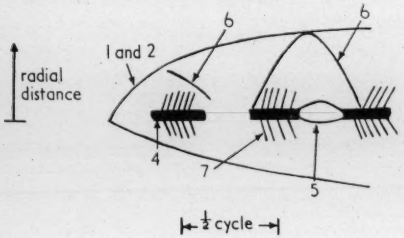


Fig. 6. Schematic streak record of a high current spark channel. The numbers refer to the description in the text

Table 2. *Data on* 188 *kA sparks in air, hydrogen and argon*

	Air	*Hydrogen*	*Argon*
Velocity of initial expansion $\times 10^{-5}$ cm/s	1·4–3·1	3·4–9·4	~2
Velocity of sound $\times 10^{-5}$ cm/s (for comparison)	0·34	1·3	0·316
Radius of channel (cm) at $t = 7{\cdot}7$ μs	1·25	2·7	~1·2
Radius of core (cm) at $t = 7{\cdot}7$ μs	0·95	0·55	~0·8

The velocity of expansion depends on the degree of asymmetry, the lowest values given in Table 2 refer to symmetrical discharges and the highest values refer to asymmetrical discharges. The data for argon are not so precise as the other data because several channels were photographed in every case. It appears, from records such as Fig. 5(*d*, *f*), that the different channels do not always develop simultaneously. Branched channels have been observed for long sparks in air and in argon, but the effect is much more pronounced for the latter gas.[22]

A typical microphotometer plot across a hydrogen spark channel, at the instant of maximum current, is given in Fig. 7 and the lateral intensity profile derived from this plot is shown in Fig. 8. The characteristic curve of the film was obtained from the maxima of microphotometer plots of a series of rotating mirror photographs taken with a step wedge. It is possible, in general, to derive the radial intensity profile from the lateral intensity profile, assuming that self-absorption is negligible. This has not been done in the present case because the analysis only applies to a symmetrical discharge column. Also, if the spectral distribution of light intensity varies across the channel, the lateral intensity profile deduced in the above way will be inaccurate.

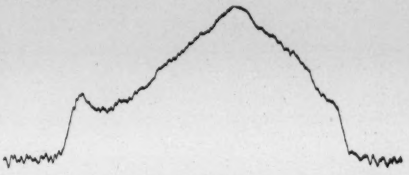


Fig. 7. A typical microphotometer plot across a 188 kA hydrogen spark channel, taken at $t = 7{\cdot}7$ μs. The scale is given in Fig. 8

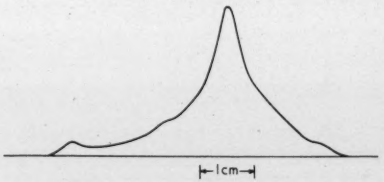


Fig. 8. The lateral intensity profile of a 188 kA hydrogen spark channel at $t = 7{\cdot}7$ μs, derived from Fig. 7

DISCUSSION

The very high velocity of the initial expansion of the channel (i.e. higher than the velocity of sound in the surrounding gas) and the luminous wavefront associated with this expansion shows that the mechanism is that of shock wave propagation.[1,19] Also, the observed asymmetrical expansion is compatible with shock wave propagation but not with diffusion, which several previous workers[1] have assumed is responsible for the expansion of the channel. The asymmetry arises in the initial breakdown of the gap.

The bright central core of the channel is probably, we suggest, due to the self-magnetic pinch effect. The theoretical papers[7–11] dealing with the pinch effect refer to steady-state, low-pressure arcs and do not apply to high-pressure (i.e. ~ 1 atm.) transient discharges. One major difference between low- and high-pressure discharges is that volume recombination is important in the latter, a statement implicit in the assumption, often made, that Saha's equation applies to high-pressure positive columns. Also, the radial variations of electron temperature, gas density and electron mobility cannot be neglected in high-pressure discharges and the electron mobility depends on electron-ion interactions in addition to electron-atom collision processes.

In spite of the complexity of the problem, certain general statements can be made. An equation of motion for the whole gas[20] can be formulated, which reduces for a steady-state, high current arc to the equation of equilibrium:

$$\frac{dP}{dr} + n_+ eE_R - n_- eE_R + jH = 0 \qquad (1)$$

where P is the total gas pressure, n_+ and n_- are the positive ion and electron concentrations, e is the electronic charge,

E_R is the radial electric field, j is the current density and H the value of the self-magnetic field at any radius r.

The second and third terms in equation (1) very nearly cancel and the equation therefore reduces to

$$\frac{dP}{dr} + jH = 0 \tag{2}$$

where H is given by

$$H = \frac{4\pi}{r}\int_0^r jrdr \tag{3}$$

If, for a rough estimate, the current density is assumed to be constant, the solution to equations (2) and (3) is

$$P = P_0 + ij\left[1 - \left(\frac{r}{r_0}\right)^2\right] \tag{4}$$

where P_0 is the ambient pressure. The total pressure at the axis exceeds that at the periphery by an amount ij dynes/cm^2, this increase in pressure can be written

$$\Delta P \simeq IJ \times 10^{-8} \text{ atmospheres} \tag{5}$$

where I and J are measured in amperes and amperes/cm^2 respectively. Equation (2) has already been stated by Mason (see Bellaschi[14]) and Alfvén.[9]

Calculated values of ΔP for the present discharges are given in Table 3. Even though the equation used for ΔP was

Table 3. *Calculated quantities referring to 188 kA sparks in air, hydrogen and argon*

	Air	*Hydrogen*	*Argon*
Current density $\times 10^{-4}$ A/cm^2	6·6	20	9·4
Power/cm^3 (MW/cm^3)	12	69	7·5
ΔP (atm.)	120	370	180
H (gauss)	40 000	68 000	47 000
Resistivity (Ω cm)	0·0027	0·0018	0·000 85
Electron temperature (°K)	72 000	94 000	160 000

derived for a steady-state discharge and the experimental results refer to transient discharges, these rough calculations show that the magnetic forces are very high in the present discharges. It is reasonable to suppose, therefore, that the bright central core of the discharge is probably due to the self-magnetic pinch effect. Considerable supporting evidence for this view is afforded by the fact that sometimes a part of the core moves out from the centre as the current decreases to zero, but returns again as the current increases [see Figs. 4(f, g) and 5(a)]. It is possible that the pinch effect may apply to transient discharges in that outward radial movements of charged particles may be inhibited in the early expanding phase of the discharges if a high self-field is produced before expansion is complete. This suggestion is consistent with the appearance of the axial streak before the rapid (shock wave) expansion process is over [see Figs. 4(f, h) and 5(a, b)], but after the elapse of some μs from the start of the spark channel phase [see, for example, Fig. 5(b)].

The inward-moving waves [item (6) in description of photographs] can be identified as shock waves. Fig. 5(f) clearly shows how the second inward wave is due to the reflexion of an outward shock wave at the channel boundary. It is not clear how the first inward moving wave arises, although it may also be due to a reflexion, if the energy input has a minimum shortly after breakdown. Assuming that the wave is a shock wave we have

$$v \geqslant \sqrt{\frac{\gamma kT}{M}}$$

where the equality sign refers to the limiting case of a weak shock, i.e. a sonic disturbance. In 188 kA hydrogen sparks, velocities of 5×10^5 cm/s have been measured. Substitution of this figure gives

$$T \leqslant 1800°\text{K}$$

assuming that the gas is atomic hydrogen and that $\gamma = \frac{5}{3}$, or

$$T \leqslant 4300°\text{K}$$

assuming that the gas is molecular hydrogen and that $\gamma = 1{\cdot}407$.

Spectroscopic examination of the discharges shows practically complete dissociation. This rough calculation indicates that the temperature in the outer channel is comparatively low.

The "jets," illustrated in Fig. 5(b), may be jets of plasma issuing from the central core at bends in the channel, where the magnetic field is weakened. These "jets" can be clearly seen in hydrogen sparks which is consistent with the above suggestion since sparks in hydrogen have a tendency to form many sharp bends.[22]

If these prominences are jets then the kinetic energy of the jet must be less than the enthalpy of the plasma in the core or, very roughly,

$$\tfrac{1}{2}(m + M)V^2 < (5/2)k(T_- + T_+)$$

where M is the mass of a proton and V is the jet velocity. Values of about $2{\cdot}5 \times 10^6$ cm/s were measured for jets occurring near the current maximum in 188 kA hydrogen sparks and substitution of this figure in the above inequality gives

$$T_- + T_+ > 22\,000°\text{K}$$

which seems reasonable.

The voltage gradients are high and increase with increasing current in the range studied (as already mentioned). The values given in Table 1 refer to the linear portion of the voltage drop-gap length curves; it appears that the gradients are higher near the electrodes than in the middle of the gap. Sufficiently long gaps were used in the rotating-mirror experiments for the channel to become uniform in the centre of the gap (according to the voltage drop-gap length curves).

The resistivity of the channel core, which is assumed to carry most of the current, has been calculated from the experimental data of Tables 1 and 2. The value so obtained was used to calculate the electron temperature assuming that

$$n_+ \bar{Q}_+ \gg N\bar{q}$$

(where N is the density of neutral atoms and $\bar{Q}_+$ and $\bar{q}$ are the mean ionic and atomic cross-sections for the transfer of momentum), i.e. assuming that the electrons in the core deliver momentum mainly to ions. This seems to be a reasonable assumption, as the pinch effect appears to be operative, so the ion concentration in the central core will be high. The data of Fig. 3 show that the longitudinal electric field strength in the high current discharges, which increases with increasing current, is greater than in sparks at lower currents ($\simeq$ 70 V/cm in hydrogen at 110 A,[19]) where the current-field strength characteristic has a negative slope. Similar differences are found between high and low current arcs. The explanation, in the spark case, presumably involves an energy loss mechanism (perhaps increased recombination

at higher electron densities in a radially restricted discharge channel) not fully operative at lower currents. The effective conduction losses from a high-temperature channel vary in a complicated fashion with temperature[21] and include the energy drain due to outwardly diffusing, charged particles which should be greatly decreased in a strong magnetic self-field.

Various estimates of $\bar{Q}_+$ have been made (Chapman and Cowling,[20] Gvosdover,[23] Davydov,[24] Alfvén[9] and others), which have a similar algebraic form, but different constants. Gvosdover's result has been used in mobility calculations by Mohler,[25] Druyvestyn and Penning,[26] Elenbaas,[27] Edels and Craggs,[28] Craig and Craggs,[19] and others. Mohler's experiments showed that the resistivity of highly ionized caesium vapour had a variation with electron temperature corresponding to that deduced from Gvosdover's expression for $\bar{Q}_+$.

If $n_+\bar{Q}_+ \gg N\bar{q}$, the resistivity depends mainly on the electron temperature and is almost independent of pressure; in fact we have

$$\rho \simeq 5{\cdot}2 \times 10^4 T_-^{-\frac{1}{2}}\ \Omega\,\text{cm}$$

The constant quoted is the one determined experimentally by Mohler. The electron temperatures derived from this equation are given in Table 3. The electron, positive ion and gas temperatures in high-pressure sparks are expected to be roughly equal, even after only a few microseconds, due to the high electron collision frequency.[1,19]

The pressures and temperatures given in Table 3 are extremely high. It is important, therefore, to examine the energy requirements and ascertain whether such pressures and temperatures are energetically possible. The case of the hydrogen spark will be considered.

About $2{\cdot}9 \times 10^{19}$ particles/cm^3 are necessary to produce a pressure of 370 atmospheres at a temperature of 94 000° K. If the temperature is roughly constant across the core, then the pressure variation is parabolic and the number of particles per centimetre associated with the increase in pressure is

$$(2{\cdot}9 \times 10^{19})2\pi\int_0^{r_0} r[1-(r/r_0)^2]dr \text{ or } \tfrac{1}{2}\pi r_0^2 \times 2{\cdot}9 \times 10^{19}$$

Assuming that most of these particles are charged, the number of ion pairs per centimetre is $\frac{1}{4}\pi r_0^2 \times 2{\cdot}9 \times 10^{19}$ or $6{\cdot}9 \times 10^{18}$ ion pairs/cm.

The dissociation energy/ion pair	= 2·2 eV
the ionization energy/ion pair	= 13·5 eV
and the thermal energy/ion pair	= 2 × 94 000/7730 = 24·3 eV
so the total energy/ion pair	= 40 eV

Thus the energy associated with the particles held in the core is $6{\cdot}9 \times 10^{18} \times 40 \times 1{\cdot}59 \times 10^{-19}$ joules, or 44 joules, which is not excessive, considering that the measured power input per centimetre is 66 J/μs. A rough estimate of the energy associated with the shock wave (following Flowers[3]) gives a value of 24 J.

Clearly much more experimental work is required, for example the direct measurement of electron density and temperature by measuring the breadth of Stark broadened Balmer lines[1] and the distribution of intensity in the Balmer series limit continuum[29] respectively. These experiments and other relevant investigations are being carried out in this laboratory by Mr. C. J. Braudo, and preliminary Stark effect observations with a spectrometer slit crossing the channel tend to confirm an axial increase in the concentration of electrons.

ACKNOWLEDGEMENTS

The authors wish to thank Professor J. M. Meek for the laboratory facilities provided. One of the authors (J. E. A.) is indebted to the University of Liverpool for the award of a Post-Graduate Studentship and later a Fellowship.

REFERENCES

(1) Meek, J. M., and Craggs, J. D. *Electrical Breakdown of Gases* (Oxford: Clarendon Press, 1953).
(2) Norinder, H., and Karsten, O. *Ark. Mat. Astr. Fys.*, **36A**, No. 16, p. 1 (1949).
(3) Flowers, J. W. *Phys. Rev.*, **64**, p. 225 (1943).
(4) Durnford, J., and Reynolds, P. I.E.E. Monograph No. 70 (1953).
(5) Allen, J. E. *Research*, **3**, p. 527 (1950).
(6) Reynolds, P., and Craggs, J. D. *Phil. Mag.*, **43**, p. 258 (1952).
(7) Bennett, W. H. *Phys. Rev.*, **45**, p. 890 (1934).
(8) Tonks, L. *Trans Electrochem. Soc.*, **72**, p. 157 (1937); *Phys. Rev.*, **56**, p. 360 (1939).
(9) Alfvén, H. *Cosmical Electrodynamics* (Oxford: Clarendon Press, 1950).
(10) Thonemann, P. C., and Cowhig, W. T. *Proc. Phys. Soc.* [*London*], *B*, **64**, p. 345 (1951).
(11) Blackman, M. *Proc. Phys. Soc.* [*London*], *B*, **64**, p. 1039 (1951).
(12) Ware, A. A. *Phil. Trans Roy. Soc.*, **243A**, p. 197 (1950–51).
(13) Cousins, S. W., and Ware, A. A. *Proc. Phys. Soc.* [*London*], *B*, **64**, p. 159 (1951).
(14) Bellaschi, P. L. *Elect. Engng*, **53**, p. 86 (1934); **54**, p. 837 (1935); *Elect. J.*, **32**, p. 237 (1935); **33**, p. 273 (1936); *Elect. Engng*, **56**, p. 1253 (1937); **58**, p. 466 (1939).
(15) McEachron, K. B., and Thomason, J. L. *Gen. Elec. Rev.*, **38**, p. 126 (1935).
(16) Stekolnikov, I. S. *C.R. Acad. Sci.* [*U.R.S.S.*], **52**, p. 413 (1946).
(17) Pfestorf, G. *Elektrotech. Z.* [*ETZ*], **73**, p. 124 (1952).
(18) Beams, J. W. *Phys. Rev.*, **35**, p. 24 (1930).
(19) Craig, R. D., and Craggs, J. D. *Proc. Phys. Soc.* [*London*], *B*, **66**, p. 500 (1953).
(20) Chapman, S., and Cowling, T. G. *The Mathematical Theory of Non-Uniform Gases* (London: Cambridge University Press, 1939).
(21) Höcker, K. H., and Schulz, P. *Z. Naturforsch.*, **4a**, p. 266 (1949).
(22) Craggs, J. D., Hopwood, W., and Meek, J. M. *J. Appl. Phys.*, **18**, p. 919 (1947).
(23) Gvosdover, S. *Phys. Z. Sowjet.*, **12**, p. 164 (1937).
(24) Davydov, B. *Phys. Z. Sowjet.*, **12**, p. 269 (1937).
(25) Mohler, F. L. *J. Res. Nat. Bur. Stand.*, **21**, p. 873 (1938).
(26) Druyvesteyn, M. J., and Penning, F. M. *Rev. Mod. Phys.*, **12**, p. 87 (1940).
(27) Elenbaas, W. *The High Pressure Mercury Vapour Discharge* (Amsterdam: North-Holland Publishing Co., 1951).
(28) Edels, H., and Craggs, J. D. *Proc. Phys. Soc.* [*London*], *A*, **64**, p. 574 (1951).
(29) Ravenhill, P., and Craggs, J. D. *Proc. Phys. Soc.* [*London*], *B*, **64**, p. 618 (1951).

Correspondence

The correction of an error in the determination of high resistances by capacitor discharge

Referring to the paper by Dr. G. C. Curtis published in the March issue of your *Journal*, I should like to point out that I dealt with a very similar problem, that of finding the leakage resistance of a long cable, more than 20 years ago.(1) Later, I extended the treatment to the analogous heat-transfer problem.(2)

I see that Dr. Curtis does not attempt the solution, but is content to rely on one given by Carslaw and Jaeger. These gentlemen are, of course, notable (or even notorious) proponents of the Laplace transform method of solving transients; but I obtained my solution in 1932 by the older (and, to me, simpler) method of Heaviside.

Royal Naval College, Greenwich. D. K. McCLEERY

REFERENCES

(1) McCLEERY, D. K. *Proc. Phys. Soc.* [*London*], **44**, p. 494 (1932).
(2) McCLEERY, D. K. *Nature* [*London*], **130**, p. 737 (1932).

In the papers he refers to, Instructor-Commander McCleery dealt with a transmission line free at both ends, lumped capacity being absent, a steady input current to which is broken at time zero. Thus his boundary conditions and mine differ so much that the problems are completely distinct.

As regards his criticism of the mathematical apparatus invoked, I think that most people nowadays would consider the transform method to be an improvement on the original Heaviside method, but this question is one of personal preference.

The remaining point seems to amount to a criticism of the fact that no proof was given. It seems to me that in a journal devoted to applied physics it is best to replace long mathematical derivations by literature references wherever possible.

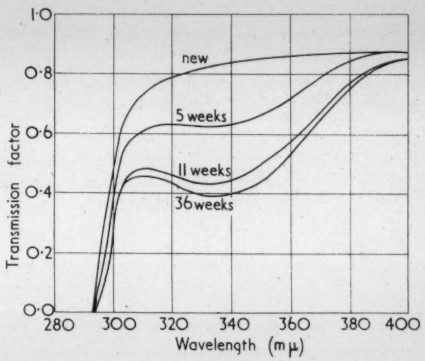
United Kingdom Atomic Energy Authority, Sellafield, Cumberland. G. C. CURTIS

The solarization of polymethylmethacrylate

The need to study the solarization of perspex arose during a programme of measurements of ultra-violet solar radiation. The photocell and first amplifying stages of the recorder which had been constructed for these measurements were sealed into a container with a perspex window. When new, the window transmitted most of the solar radiation to which the photocell was sensitive (wavelengths 290 to 330 mμ), but a decline in the readings of the recorder in the first few months of use indicated that the transmission of the perspex decreased under the action of ultra-violet solar radiation.

Three samples of the perspex used for the window of the photocell container were therefore placed outside in the sun, and their spectral transmission was measured in the Laboratory at regular intervals. The measurements were made by the normal spectrophotometric methods, based on the use of a single-prism quartz spectrometer, with a filter to reduce stray light. The source was a mercury–cadmium discharge lamp, and measurements were made at the wavelengths of fifteen spectral lines in the range 290 to 400 mμ. Surface dirt was rubbed off the samples before each measurement.

The result of the exposure to the sun, as shown in the figure, was a progressive decrease in the transmission factor at all



The progressive decrease in ultra-violet transmission of polymethylmethacrylate (plasticized perspex No. 5, thickness 1·42 mm) on exposure to the sun

ultra-violet wavelengths. The greatest effect was in the band 320 to 360 mμ. No absorption band of this type appeared in the glasses and other window materials studied by Coblentz and Stair,(1, 2) but it is very similar to the one produced in polystyrene in the experiments of Reiney, Tryon and Achhammer.(3)

It should be noted that these measurements were made mostly during the spring, when the ultra-violet solar radiation is at a high level; at other times of the year the solarization would be less rapid.

The work described in this note was carried out in the Dominion Physical Laboratory of the New Zealand Department of Scientific and Industrial Research, and is published with the permission of the Secretary of the Department.

Dominion Physical Laboratory, Department of Scientific and Industrial Research, New Zealand. K. J. McCREE

REFERENCES

(1) COBLENTZ, W. W., and STAIR, R. *J. Res. Nat. Bur. Stand. Wash.*, **3**, p. 629 (1929).
(2) COBLENTZ, W. W., and STAIR, R. *J. Res. Nat. Bur. Stand. Wash.*, **13**, p. 773 (1934).
(3) REINEY, M. J., TRYON, M., and ACHHAMMER, B. G. *J. Res. Nat. Bur. Stand. Wash.*, **51**, p. 155 (1953).

The effect of intensifying screens on X-ray line breadths

When line positions in X-ray diffraction work are required, intensifying screens are often used to reduce exposure times, particularly when crystal reflected radiation, or X-radiation of short wavelength such as from molybdenum, is employed. It is well known that such screens can introduce errors in intensity measurements since the intensification produced depends upon such factors as the type of screen, type of film,

film density, X-ray wavelength and intensity.[1] It is, however, perhaps not generally appreciated how serious are the errors that they may introduce in line breadth determinations.

In the course of work, described elsewhere,[2,3] on the X-ray line broadening from cold-worked iron, a few subsidiary experiments were carried out to determine the importance of this effect. Back reflexion patterns were obtained with filtered molybdenum radiation from two iron specimens, one annealed and the other plastically deformed. In one set of exposures a non-screen film (type New Ilfex by Ilford Ltd.) was used with no intensifying screen. Duplicate exposures were made on a screen type film [type Blue Brand by Kodak (Australasia) Pty. Ltd.] with a Fluorazure intensifying screen (made by Ilford Ltd.). In each case a 0·005 in. aluminium filter was placed directly in front of the film to reduce the scattered radiation. The integral line breadths, b and B, obtained by microphotometry of the lines, from the annealed and deformed specimens respectively are given below. The intensifying factor due to the screen for the peak intensity was approximately 7·5.

If, for purposes of comparison, the true breadth β is taken to be the difference between B and b, it will be seen that the use of the intensifying screen has reduced β to less than half its true value. A calcium tungstate screen reduced β by only about 3%, but in this case the peak intensifying factor was 1·15.

Integral line breadths ($\times 10^2$) *in radians for* 651, 732 *reflexions*

Film and screen	*b*	*B*	*B-b*
Ilford New Ilfex with no screen	1·23	2·76	1·53
Kodak Blue Brand with Fluorazure screen	1·21	1·93	0·72

These results are, of course, only representative of one particular set of experimental conditions, but they serve to illustrate that completely misleading results can be obtained from a combination of short wavelength radiation, a high screen intensifying factor and wide X-ray lines.

Aeronautical Research Laboratories, Melbourne, Victoria, Australia.
R. I. GARROD
J. H. AULD

REFERENCES

(1) GAMERTSFELDER, C., and GINGRICH, N. S. *Rev. Sci. Instrum.*, **9**, p. 154 (1938).
(2) AULD, J. H., and GARROD, R. I. *Nature* [*London*], **169**, p. 579 (1952).
(3) GARROD, R. I., and AULD, J. H. *Acta Met.*, to be published.

New books

Progress in biophysics and biophysical chemistry, Vol. 4. Edited by J. A. V. BUTLER and J. T. RANDALL, F.R.S. (London: Pergamon Press Ltd.). Pp. viii + 339. Price 63*s*.

This volume reaches the same high standards as the first three volumes and will be a welcome addition to the libraries of teachers and research workers in the fields with which it deals.

It contains seven separate reviews of topics which have quite close relationships with one another. The first, on "Polyelectrolyte gels" by A. Katchalsky, gives a factual account in which attention is devoted primarily to the simpler biogels and to gels of synthetic polyelectrolytes. The second article, by H. H. Weber and Hildegard Portzehl discusses "The transference of the muscle energy in the contraction cycle." The facts which appear to be practically certain are reviewed and an attempt is made to connect the working cycle with the events of excitation. An excellent account of the working cycle of the living muscle is given at the end as a summary of the observations previously described. H. Fernandez-Morán has reviewed "The submicroscopic structure of nerve fibres," dealing mainly with the results obtained so far in examinations of peripheral and central myelinated nerve fibres and of certain types of unmyelinated fibres by means of the electron microscope. As a necessary background, other studies with the light microscope, polarized light, X-ray diffraction and other physical methods are summarized. "The nucleoprotein complex of the cell nucleus and its reactions" is the title of the fourth article by P. F. Davison, B. E. Conway and J. A. V. Butler which gives a descriptive account of the results to date and theories of the observations. P. C. Koller has reviewed the subject of "Chromosome breakage". In this review criticism is levelled at the mathematical theory of breakage by radiation in accordance with the target theory and new evidence is described in support of this criticism. It is emphasized that chromosome breakage should be regarded as an essential part of the general biochemical mechanism underlying cell reaction, adaptation and selection. The sixth review, by J. C. Kendrew, is of "The crystalline proteins: recent X-ray studies and structural hypotheses" in which the author attempts to satisfy the interests of biologists without being too optimistic about present knowledge. The whole review stresses that the tree of protein crystallography is not a prolific bearer of prize fruits but, nevertheless, no other techniques offer hope of solving the kind of problem that can be solved by X-ray crystallography. The final article, "Facts and theories about muscle" by Dr. D. R. Wilkie summarizes the properties of living muscle and describes and analyses several theories of contraction.

All the articles are well written, readable and thoroughly documented. From the student's viewpoint perhaps the only thing against all four of the volumes in this series is their price. This surely prevents many from purchasing who otherwise would prefer their own copy. C. W. WILSON

Heat conduction, with engineering, geological and other applications. By L. R. INGERSOLL, O. J. ZOBEL and A. C. INGERSOLL. (Madison: The University of Wisconsin Press, 1954). Pp. xiii + 325. Price $5.00.

Just over 40 years ago Ingersoll and Zobel wrote a slim volume of 172 pages on the subject of mathematical theory of heat conduction. To successive generations of students the book has been of the greatest assistance particularly to those who had neither the time nor the mathematical preparation to read Carslaw's *Mathematical theory of the conduction of heat in solids*.

The volume under review is the much enlarged version of the earlier editions and is the joint work of three authors, L. R. Ingersoll, O. J. Zobel and A. C. Ingersoll. It runs to 325 pages and brings the subject up to date. Noteworthy additions are those relating to the theory of earth heat exchangers for the heat pump and drying and soil consolidation.

As indicated in the title engineering, geological, and other applications occupy an important place. A feature of particular importance to those whose interests are largely on

the practical side is the discussion of auxiliary graphical and other approximation methods by which many practical heat conduction problems may be solved with only the simplest mathematics. The only mathematical prerequisite necessary for reading the book is a reasonable knowledge of the calculus.

Another valuable feature of the book to the student is the set of problems at the end of each chapter. Two hundred references are given to the literature of the subject. The authors are to be congratulated on the production of a book which will be helpful to a wide circle of readers.

E. GRIFFITHS

Proceedings of the second international congress on rheology, Oxford, 26–31 July, 1953. Edited by V. G. W. HARRISON, Ph.D., F.Inst.P. (London: Butterworth's Scientific Publications; New York: Academic Press Ltd.) Pp. ix + 451. Price 60*s*.

The appearance of this volume marks an important stage in the development of the science of rheology. When one compares it with such early classics as the *First report on viscosity and plasticity* (1935), one is made aware of the enormous expansion which has taken place during the last twenty years. This expansion has been due very largely to the industrial development of a great variety of new materials, among which may be mentioned in particular materials of a high-molecular character, whose mechanical properties cannot be represented in any simple manner. For the study of these materials a rheological outlook is essential.

The present volume includes all the papers presented at the Congress, together with the discussions. The papers are arranged under the following sectional headings: high polymers; viscosity and plasticity; biology; oils and greases. In addition there are six general lectures and the address delivered by the President, Sir Geoffrey Taylor.

The papers generally are of a high quality, and the complete work is a valuable record of a highly successful international enterprise. It will be welcomed not only by those who attended the Congress but by a much wider circle of workers in rheology and allied fields. The publishers deserve credit for an excellent production at a moderate price, and the Editor especially is to be congratulated for making it possible.

L. R. G. TRELOAR

Geometrical mechanics and de Broglie waves. By PROF. J. L. SYNGE., Sc.D., F.R.S. (London: Cambridge University Press, 1954.) Pp. vi + 167. Price 25*s*.

Hamilton's method in optics postulates only the existence of a medium function, giving the refractive index as a function of position (and of direction in an anisotropic medium). The calculus of variations does the rest; rays being associated with extremals, and wave surfaces with transversals. Physically, of course, the method often becomes inapplicable, for example, in the neighbourhood of sources and foci, but mathematically it remains valid. If Newtonian mechanics is formulated in terms of a medium function, a formally identical theory results, particle trajectories are the counterparts of rays, and surfaces of constant action correspond to wave surfaces. The primary task that Prof. Synge has set himself is to complete this theory of "geometrical mechanics" by applying Hamilton's optical method to Minkowski space-time. By adding to this formal theory de Broglie waves it is possible to treat problems as "physical mechanics." In this way the theory of the hydrogen atom is treated, and it leads to the correct fine structure and Zeeman effect formulae.

The level of treatment is described as "essentially simple," but that description would not be readily accepted by a reader lacking more than a nodding acquaintance with the kind of language used in the quotation (p. 12) "the 3-surface of slowness W is the polar reciprocal of the unit 3-wave S (and, of course, conversely) with respect to the unit pseudosphere $y_r y_r = -1$, that is, the unit pseudosphere lying inside the null-cone." However, anyone with anything of a taste for mathematical physics will not be deterred by this.

The natural field of application of these methods would seem to be that of the optics (or mechanics, if you will) of fast-moving free electrons, and it is to be expected that further applications of this kind will be made.

H. H. HOPKINS

The amplification and distribution of sound. 3rd edition, revised. By A. E. GREENLEES, A.M.I.E.E. (London: Chapman and Hall Ltd., 1954.) Pp. x + 300. Price 35*s*.

The earlier editions of this book were very favourably received by the scientific press and it is not surprising, therefore, that a third edition should have been found necessary. The author has taken advantage of this to revise his text and to add sections on items which have undergone technical development since the war. The treatment is kept simple by avoiding elaborate mathematics and including, at many points, numerical examples, so that the book is well suited to those engineers and students who have only a meagre mathematical equipment. Furthermore, a chapter on fundamentals clarifies the meaning and use of electrical and acoustical terminology; indeed, it is not even assumed that the reader knows Ohm's law though he is soon involved in quite elaborate circuitry.

A large part of the text is devoted to a comparative description of types of equipment—amplifiers, microphones, loudspeakers, recorders and reproducers—and the means of making measurements thereon. The only notable omission is the electrophonic organ, but its equipment is perhaps too elaborate for a book of this nature. The chapter on installation planning forms an elementary introduction to acoustics of buildings.

The book can be recommended to all who have to deal with electro-acoustic equipment. E. G. RICHARDSON

Atomphysik, Volumes III and IV. By Dr. KARL BECHERT. (Berlin: Walter de Gruyter and Co., 1954.) Pp. 148 and 170. Price *DM* 4.80 each.

The Sammlung Göschen has long been famous for its brief, authoritative monographs in almost every field of knowledge. This new series of seven volumes on atomic physics promises to be most valuable to students and perhaps even more to teachers. Only Volumes III and IV are available as yet. Volume III begins with a careful exposition of the uncertainty relations; after a discussion of Bohr's atomic model there follow chapters on quantum mechanics, giving the formalisms of wave mechanics, matrices and operators, and ending with the (non-relativistic) treatment of the hydrogen atom. Volume IV deals with atoms with two and more electrons, perturbation theory, spin (including Dirac's theory of the electron), and molecules.

The style is very compact and precise; no loose ends, no ambiguities for the careful reader. The mathematical tools of atomic theory are defined and used as the need arises. The factual style and crammed information recalls a travel guide book. Not an easy book for a student to tackle, but very rewarding to those who have the courage, and admirable for helping a teacher to prepare his course.

O. R. FRISCH

Amplitude-frequency characteristics of ladder networks. By E. GREEN, M.Sc. (Chelmsford: Marconi's Wireless Telegraph Co. Ltd., 1954.) Pp. iii + 155. Price 25*s*.

In the first part of this work the author discusses the general properties of low-pass ladder networks, their response characteristics and application to broad-band networks. Design data are expressed graphically and in tabular form for the several coefficients in the equations. The second part deals with several applications of these networks to such functions as the coupling between valves and resistive load, broad-band matching of reactive loads, low-pass and band-pass filters, design of networks to obtain optimum performance of valves.

The book should find its main use as a work of reference for qualified telecommunication engineers and not as a treatise for students. It is pleasant to find a specialized work of this nature so reasonably priced. A. J. MADDOCK

Reports on progress in physics, Vol. XVII, 1954. Edited by A. C. STICKLAND. (London: The Physical Society, 1954.) Pp. 280. Price 50*s*.

Volume XVII maintains the high standard set by previous volumes. It contains eight articles each in the form of a model monograph written by an expert on the topic concerned. As usual the volume covers a wide field and both theoretical and experimental advances are represented.

The longest and most detailed article is by C. J. Bouwkamp and surveys recent progress in classical diffraction theory. The impetus to new developments in this field has come from radio microwave technique and it appears that since 1940 an almost overwhelming number of papers (over 500 references are given) on diffraction theory have been published. The next longest article is on antiferromagnetism by A. B. Lidiard; one very useful feature here is a table listing experimental papers on antiferromagnetism according to the compound and the property studied. There are two reports on atmospheric physics, one by J. A. Chalmers on atmospheric electricity and the other by Briggs and Spencer on horizontal movements in the ionosphere. Nuclear physics is represented by a theoretical article by M. H. L. Pryce on nuclear shell structure. Another theoretical article is contributed by W. Moffitt and deals with atomic valence states and chemical binding. Lastly there are two shorter articles on topics which are likely to be of more general interest than the above; in one C. W. Allen considers the solar corona and in the other E. Teller discusses a theory of the origin of cosmic rays.

It would be unfair in so short a review to select any article for special comment. All are well written and well provided with references and each gives an abstract so that the scope can be seen at a glance. P. NICOLSON

Structure and properties of solid surfaces. Edited by Robert GOMER and CYRIL S. SMITH. (Chicago: Chicago University Press; London: Cambridge University Press.) Pp. xvi + 491. Price 64*s*.

This compilation consists of fourteen papers read (and discussions following) at a conference arranged by the National Research Council, Lake Geneva, Wisconsin, in September 1952. The foreword is dated 1953. It is only necessary to list the authors' names to realize that this report is clearly of the utmost value to all those engaged in the study of solid surfaces. Papers have been contributed by C. Herring, P. P. Ewald and H. Juretschke, A. J. Shaler, W. A. Weyl, G. P. Thomson, F. P. Bowden and D. Tabor, A. F. Wells, H. E. Buckley, H. Seifert, T. L. Hill, M. Boudart, P. H. Emmett, A. Wheeler and G. M. Schwab, E. Roth, C. Grintzos and N. Mavrakis. The aspects dealt with range from theories of surface energy through theories of adhesion, crystal growth, epitaxy, sorption and catalysis. I am convinced that workers in every branch of physics, chemistry and metallurgy dealing with surfaces of solids will find a great deal of personal interest and very real value in almost every paper, no matter what his field of interest, so comprehensive is the range and so authoritative are the discussions. The papers are mainly in the nature of reviews and this is the particular value of the book since in its 490 pages it brings one up to date in so many different modes of approach. I have enjoyed every page I have read and I feel sure that this will be the opinion of many. Primarily the book is of value to the research worker, but since the sections are self-contained, even an undergraduate, in his final year, may glean much of value from individual sections. Although the book is expensive it is well worth the price and should be at hand for every research worker engaged in studies on solid surfaces.
S. TOLANSKY

Technische hydro- und aeromechanik. By WALTHER KAUFMANN, Dr.-Ing.habil. (Berlin: Springer-Verlag, 1954.) Pp. viii + 352. Price *DM*36.

The amount of information contained within the 352 pages of this book is surprisingly large. It ranges from hydrostatics to compressible flow and on the way takes in one-dimensional flow; incompressible flow including 2 and 3 dimensional aerofoil theory and waves; viscous motion including slow motion, lubrication and boundary layer theory; the stability of viscous flow; turbulence, including the mixing hypotheses and boundary layer theory.

It is not surprising, therefore, that many of the topics get rather brief treatment. Those your reviewer found particularly tantalizing were the treatments of separation of the boundary layer, the stability of viscous flow and the section on compressible flow. In the last-named the general equations of motion are given without apparently any mention of the corresponding general energy equation. It is very difficult to see how the former can be used without the latter except in special cases where the complexities of the former are no longer necessary.

One is left with the impression that the book is written for those who want to know the answers without being too concerned about the fundamental reasons. Viewed from this aspect the book can be recommended as a mine of useful information whose value in the reviewer's opinion, suffers to some extent because the references, though numerous, are not even more extensive and because an author index is omitted. L. HOWARTH

A rapid analytical method for the determination of alumina, silica and iron oxide in refractories by the photometry of organic complex reagents. By K. GIESEN, O. GLEMSER and E. RAULF. **Investigations of the alkali-content of refractories using the Riehm-Lange flame photometer.** By K. GIESEN and P. KAMPA. No. 59 of the Forschungsberichte des Wirtschafts und Verkehrsministeriums Nordrhein Westfalen. (Westdeutscher Verlag, Koln u. Opladen, 1954.) Pp. 51. Price *DM*11.60.

These two analytical methods, which are complementary to each other in the rapid analysis of refractories, are each described in great detail in this booklet. Useful data on the stability of certain coloured complexes and on the inter-element interferences which complicate the analysis of most refractories and minerals are included. The two methods form part of a series of progress reports edited by Professor Leo Brandt. E. VAN SOMEREN

Notes and comments

Elections to The Institute of Physics

The following elections have been made by the Board of The Institute of Physics:

Fellows: J. M. Brooks, G. Harries, I. Maddock, G. R. Newbery, G. W. Sutton, J. Vennart.

Associates: M. A. Cayless, G. Eaves, G. Evans, M. H. Francombe, F. N. Goode, G. D. R. Granick, R. D. Harris, S. T. Hermiston, A. Hooper, H. Kaiser, J. R. A. Lakey, H. W. Loeb, J. G. McCormack, H. R. Milne, G. C. E. Olds, P. Howarth.

Thirty-three Graduates, thirty-five Students and seven Subscribers were also elected.

Recent advances in meteorology—Toronto Conference, 1953

In September 1953 in Toronto some 200 meteorologists attended the first large scale joint meeting of the American Meteorological Society and the Royal Meteorological Society. Their activities were arranged in the form of symposia on various branches of meteorology and the papers presented have now been published by the Royal Meteorological Society.

The subjects range widely in scale and application and include ozone distribution, arctic meteorology, climatic change, micrometeorology and cloud physics. To give some examples of individual contributions, one paper deals with the use of rocket soundings over New Mexico in studying ozone distribution up to a height of 70 km, while at the other end of the height scale there is a paper dealing with the investigation of the longer cores of material which can now be extracted from the sea bed in search of evidence of climatic change. Another paper considers the practical difficulties of setting up and maintaining weather stations in the Arctic, while in yet another it is interesting to note that use has been made of aircraft exhaust trails in studying oscillatory and turbulent disturbance in the atmosphere. Finally, since no collection of papers on applied meteorology would be complete without some reference to atmospheric pollution, it is noteworthy that the present series contains one in which the bearings of pollution on climate and health are discussed.

The papers mentioned are but a few of about fifty interesting and stimulating papers in this volume which provide a valuable record of the present progress and trends in the broad field of meteorological science.

The volume can be obtained from the Royal Meteorological Society, 49 Cromwell Road, London, S.W.7, price 30*s.* (including postage).

Erratum

In the paper *The use of semiconductors in thermoelectric refrigeration*, by H. J. Goldsmid and R. W. Douglas, on page 386 of the November issue of this *Journal*, the scale of the ordinate in Fig. 3 should be multiplied by a factor of 2. The correct optimum values of ϵ/kT for $(I_n + I_p)/k = 8$, 5 and 4 respectively are $-0{\cdot}84$, 0 and $0{\cdot}35$. The optimum thermo-electric powers in equation (10) should therefore be:

$\eta = k(4 - 0{\cdot}84) = 273\ \mu\text{V.}\,^\circ\text{C}^{-1}$ for ionized impurity scattering.

$\eta = k(2{\cdot}5) \quad = 216\ \mu\text{V.}\,^\circ\text{C}^{-1}$ for polar-type lattice scattering.

$\eta = k(2 + 0{\cdot}35) = 203\ \mu\text{V.}\,^\circ\text{C}^{-1}$ for covalent-type lattice scattering.

British Journal of Applied Physics

Original contributions accepted for publication in future issues of this Journal

A method of identifying double-flash exposures. By N. Dombrowski.
Some properties of a simple omegatron-type mass spectrometer. By A. G. Edwards.
The preparation of uniform plastic films. By R. S. M. Revell and A. W. Agar.
The measurement of ionization methods of the peak kilovoltage across X-ray tubes. By J. R. Greening.
Single exposure photography of a high speed event. By K. R. Tuson.
On the laws of creep and stress relaxation in solids. By P. Feltham.
Phase and amplitude balance methods for permittivity measurements between 4 and 50 cm. By T. J. Buchanan and E. H. Grant.
Formulae for the transformation of indices in twinning and some data for a cubic case. By K. W. Andrews and W. Johnson.
Some dynamic mechanical properties of polyisobutylene over a wide temperature range. By D. A. Thomas and D. W. Robinson.
An electrical analogue of creep and recovery: its physical and computing significance. By A. J. Kennedy.
Estimation of specific surface of the soil from mechanical analysis data. By J. R. H. Coutts.
A method of measuring the strengths of radio-frequency electric and magnetic fields in resonant cavities. By J. G. Linhart and T. H. B. Baker.

Journal of Scientific Instruments

Contents of the December issue

ORIGINAL CONTRIBUTIONS

Papers

Sensitivity of a vibrating reed, null indicator. By I. J. Somervaille.
A portable Van de Graaff generator. By T. R. Foord.
A simple circuit for use with condenser strain gauges. By D. M. Maurice.
Thermocouples suitable for use at low temperature. By W. B. Pearson.
A direct reading instrument for the measurement of small displacements. By W. D. Corner and G. H. Hunt.
The construction and calibration of a Bunsen ice-calorimeter. By L. E. Leake and E. T. Turkdogan.
A 1–100 Ω build-up resistor for the calibration of standard resistors. By B. V. Hamon.
The transient response of direct current amplifier systems. By J. H. Sanders.
A method for constructing direct reading thermistor thermometers. By J. A. McLean.
A simple microscope attachment for observing high-temperature phenomena. By J. H. Welch.
Low energy measurements with the photomultiplier scintillation counter. By G. T. Wright.
Dielectric test cell for bituminous material. By C. Q. Keiller.
A sensitive miniature pressure cell. By S. C. Redshaw.
A physical colour temperature comparator. By W. J. Brown.

Laboratory and workshop notes

A simple electromanometer. By M. Ainsworth and J. W. Eveleigh.
Gas-seal for high-temperature extensometer tensile tests in inert atmospheres. By A. J. Fenner and G. Willoughby.
Simplified construction of a straight-through metal, O-ring-sealed, vacuum stopcock. By G. W. Green.
Concurrent use of several electrode assemblies with a standard pH meter. By A. W. Kenchington and A. G. Ward.
An automatic liquid nitrogen level regulator for cold traps. By M. Feld and F. S. Klein.

NOTES AND NEWS

Correspondence

The transit time of electrons in photomultipliers. From E. H. Rhoderick and R. W. P. McWhirter.
A method for recording logarithmic variations of resistance. From H. A. Vodden.
Range finder without moving mechanical parts. From E. W. H. Selwyn.
New books
Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics, 47 Belgrave Square, London, S.W.1. (*Telephone:* Sloane 9806.) Prospective authors are invited to prepare their scripts in accordance with the *Notes on the preparation of contributions.* (Price 2*s.* 6*d.* including postage.)

REPRODUCTION. The Institute of Physics is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. Walter Judd Ltd., 47 Gresham Street, London, E.C.2. (*Telephone:* Monarch 7644.)

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

SUBSCRIPTION RATES. A new volume commences each January. The charge is £4 per volume ($11.50 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 8*s.* each ($1.15 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics, 47 Belgrave Square, London, S.W.1, or to any bookseller.